

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 10:42:18 ; Search time 342.5 seconds
(without alignments)

11066.024 Million cell updates/sec

Title: US-09-912-559-1

Sequence: 1 atgtttccagatgtctga.....aaagtgaagtgcttctaa 1683

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: N.Geneseq_101002.*
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1683	100.0	1683	24	AA145696
2	1683	100.0	3008	24	ABN95170
3	1679.8	99.8	1683	24	AA145697
4	1678.2	99.7	3623	21	AACT6693
5	526.4	31.3	617	21	AA44763
6	242.6	14.4	397	24	AB199281
c 7	120.2	7.1	451	22	ABA58323
c 8	120.2	7.1	451	22	AAK06420
c 9	120.2	7.1	451	22	AAK32096
					Human blood coagu
					Gene #1668 used t
					Human blood coagu
					Human OPRF ORF224
					Human secreted ex
					Mouse ischemic c
					Human foetal liver
					Human brain expt
					Human bone marrow

c	10	120.2	7.1	451	22	AA137947	Probe #6633 used t
c	11	120.2	7.1	451	24	ABS06863	Human genome-deriv
c	12	117.6	7.0	2033	24	AA063951	Hepatocyte growth
c	13	117.6	7.0	2036	24	ABN95114	Gene #1612 used to
c	14	117.6	7.0	2036	24	AB165242	lung cancer relate
c	15	117	7.0	117	22	ABA70919	Human foetal liver
c	16	117	7.0	117	22	AAK19188	Human brain expres
c	17	117	7.0	117	22	AAK45153	Human bone marrow
c	18	117	7.0	117	22	AA151106	Probe #19792 used
c	19	117	7.0	117	24	ABS19403	Human genome-deriv
c	20	96.2	5.7	970	15	AA063945	Gene comprising HG
	21	90.6	5.4	2512	12	AA012867	JM1-229 cell line
	22	87.6	5.2	2252	11	AA005673	v-PA.alpha1. Desm
	23	86.8	5.2	1323	8	AAAT0892	Modified tissue pl
	24	85.2	5.1	1422	8	AAAT0891	v-PA.beta. Desm
	25	84.8	5.0	2137	11	AA005675	v-PA.alpha2. Desm
	26	79.6	4.7	2257	11	AA005674	Chicken urokinase
	27	78.8	4.7	919	18	AAAT90048	Mutated CDNA codin
	28	78.6	4.7	1689	10	AAAN91217	Vampire bat plasmi
	29	78	4.6	1620	11	AA000543	T-PA with -ve char
	30	77	4.6	2097	12	AAO12071	T-PA variant havin
	31	77	4.6	2100	12	AAO12074	tPA-2 CDNA. Synth
	32	75.4	4.5	1047	14	AAO40658	mt-PA6-E. Homo sa
	33	75.4	4.5	1065	18	AAAT87456	DNA encoding t-PA
	34	75.4	4.5	1065	18	AAAT62606	Human truncated tp
	35	75.4	4.5	1068	14	AAO53320	Human truncated tp
	36	75.4	4.5	1068	14	AAO53322	tPA-1 CDNA. Synth
	37	75.4	4.5	1068	14	AAO40657	tPA-6 CDNA. Synth
	38	75.4	4.5	1068	14	AAO40660	tPA-11 CDNA. Synth
	39	75.4	4.5	1068	14	AAO40667	tPA-12 CDNA. Synth
	40	75.4	4.5	1068	14	AAO40669	tPA-13 CDNA. Synth
	41	75.4	4.5	1068	14	AAO40671	tPA-14 CDNA. Synth
	42	75.4	4.5	1068	14	AAO40673	tPA-16 CDNA. Synth
	43	75.4	4.5	1068	14	AAO40677	tPA-17 CDNA. Synth
	44	75.4	4.5	1068	14	AAO40679	tPA-18 CDNA. Synth
	45	75.4	4.5	1068	14	AAO40681	

ALIGNMENTS

RESULT 1	AA145696	standard: DNA: 1683 BP.
ID	AA145696	
XX	AA145696:	
AC		
XX		
DT	13-JUN-2002	(first entry)
XX		
DE	Human blood coagulation factor VII activating protease DNA.	
XX		
KW	Human: blood coagulation factor VII activating protease: FSAP.	
KW	single-chain plasminogen activator; bleeding disorder; haematological;	
KW	haemostatic; gene; ds.	
XX		
OS	Homo sapiens.	
XX		
FT	Key	Location/Qualifiers
FT	CDS	1..1683
FT		/*tag= a
FT		/product= "FSAP"
XX		
PN	EP1182258-A1.	
PD	27-FEB-2002.	
XX		
PF	05-JUL-2001: 2001EP-0115691.	
XX		
PR	26-JUL-2000: 2000DE-1036641.	
PR	10-OCT-2000: 2000DE-1050040.	
PR	21-OCT-2000: 2000DE-1052319.	
PR	12-APR-2001: 2001DE-1018706.	
XX		

PA (AVER) AVENTIS BEHRING GMBH.
XX Roemisch J, Stoehr H, Feussner A, Lang W, Weimer T, Becker M;
PI Nerlich C, Muth-Naumann G;
XX WPI: 2002-270939/32.
DR P-PSDB; AA017144.
XX New nucleic acid encoding mutant factor 7 activating protease, useful
PT for diagnosis, treatment and prevention of coagulation disorders, also
PI related protein and antibodies -
XX
PS Claim 2; Page 15-16; 27pp; German.
XX The present invention relates to a mutant of the DNA sequence encoding
CC the proenzyme (FSAP) that activates blood coagulation factor VII (FVII)
CC and single-chain plasminogen activator, where at least one of the base
CC changes G to C at nucleotide 1177 and G to A at nucleotide 1601 is
CC present. The mutant sequences can be used in the treatment and prevention
CC of bleeding disorders associated with inherited or acquired defects of
CC blood coagulation factors V, VIII, IX, X, XI, XIII, von Willebrand's
CC factor and/or with antibodies against any of these proteins. The present
CC sequence is the human FSAP coding sequence.
XX
SQ Sequence 1683 BP: 440 A; 436 C; 437 G; 370 T; 0 other;
Query Match 100.0%; Score 1683; DB 24; Length 1683;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1683; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTTGGCAGATGTGATGATCCATGTCCTGCTTAATGAGTGGTGGAAAGACA 60
DB 1 ATGTTGGCAGATGTGATGATCCATGTCCTGCTTAATGAGTGGTGGAAAGACA 60
QY 61 GCCTGGGTTCCCTGATGCTTTATTTGGAAGCTGGACCAGATGAGCCCTGAC 120
DB 61 GCCTGGGTTCCCTGATGCTTTATTTGGAAGCTGGACCAGATGAGCCCTGAC 120
QY 121 CAGATGATTAAGTACGATGAGATTTAATCAGAAAGAACACCATATACACTTACC 180
DB 121 CAGATGATTAAGTACGATGAGATTTAATCAGAAAGAACACCATATACACTTACC 180
QY 121 CAGATGATTAAGTACGATGAGATTTAATCAGAAAGAACACCATATACACTTACC 180
DB 121 CAGATGATTAAGTACGATGAGATTTAATCAGAAAGAACACCATATACACTTACC 180
QY 181 CATGCTGAGATTCCTGACTGCTACTGACAGACCAAGCTGATCCATCCAGCCCAAC 240
DB 181 CATGCTGAGATTCCTGACTGCTACTGACAGACCAAGCTGATCCATCCAGCCCAAC 240
QY 241 CCTGTGAACACGGTGGGAGCTGCTGTCATGGAGACACTTCACATGCACTGCTG 300
DB 241 CCTGTGAACACGGTGGGAGCTGCTGTCATGGAGACACTTCACATGCACTGCTG 300
QY 301 GCTTCCTTTCTGCGAATAAAGTCCAGAAAGTCAAAATACGTCAAGGACAACCTATGT 360
DB 301 GCTTCCTTTCTGCGAATAAAGTCCAGAAAGTCAAAATACGTCAAGGACAACCTATGT 360
QY 361 GGGCGGGGCAATGTCATTATCCAGAGTCCCTCTACTACCCCTGTGCTTTAAACAC 420
DB 361 GGGCGGGGCAATGTCATTATCCAGAGTCCCTCTACTACCCCTGTGCTTTAAACAC 420
QY 421 CATTACAGAGTCCAGAGTCCCTCAAGAGTCCCTGATGAGGCAAAACCCCTGACAG 480
DB 421 CATTACAGAGTCCAGAGTCCCTCAAGAGTCCCTGATGAGGCAAAACCCCTGACAG 480
QY 481 AATGGGGCTACCTGCTCCCGGCATTAAGCGAGATCCAAATTCACCTGTGCTGCCGAC 540
DB 481 AATGGGGCTACCTGCTCCCGGCATTAAGCGAGATCCAAATTCACCTGTGCTGCCGAC 540
QY 541 CAGTTAAAGGGAATTCGTGAATAAGTTCGATGATGCTATGTTGGCGATGCTTAC 600
DB 541 CAGTTAAAGGGAATTCGTGAATAAGTTCGATGATGCTATGTTGGCGATGCTTAC 600
QY 601 TCTTACCGAGGAAATGATATAGACAGTCAACAGACATGCGCTTTACTGGAAGTCC 660
DB 601 TCTTACCGAGGAAATGATATAGACAGTCAACAGACATGCGCTTTACTGGAAGTCC 660

QY 661 CACCTCCTTTGCGAGAGAAATTACAAACATGTTATAGAGATGCTGAAACCCATGGATT 720
DB 661 CACCTCCTTTGCGAGAGAAATTACAAACATGTTATAGAGATGCTGAAACCCATGGATT 720
QY 721 GGGGAACACAAATTTCTGACAGAAACCCAGATGGCGAGCAAAAGCCCTGTGCTTTATATA 780
DB 721 GGGGAACACAAATTTCTGACAGAAACCCAGATGGCGAGCAAAAGCCCTGTGCTTTATATA 780
QY 781 GTTACCAATGACAAAGTGAATGGAATGGAATCTGATGCTGCTGCTGCTGCTGCTGCTGCT 840
DB 781 GTTACCAATGACAAAGTGAATGGAATGGAATCTGATGCTGCTGCTGCTGCTGCTGCTGCT 840
QY 841 GTTCCCTACCCAGAGAAAGCCCTACTGAGCCATCAACCAAGCTTCCGGGTTTACCTCC 900
DB 841 GTTCCCTACCCAGAGAAAGCCCTACTGAGCCATCAACCAAGCTTCCGGGTTTACCTCC 900
QY 901 TGTGAAAGACTGAGATAGCAGAGAGAGATCAAGAGATCTATGAGAGCTTTTAAAGAC 960
DB 901 TGTGAAAGACTGAGATAGCAGAGAGAGATCAAGAGATCTATGAGAGCTTTTAAAGAC 960
QY 961 ACGGCGGCAAGCCCATGAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB 961 ACGGCGGCAAGCCCATGAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
QY 1021 ATGCCCAAGGCACTTCTGTGTGGGGCGCTGATCCACCCTGCTGAGCTGCTGCTGCT 1080
DB 1021 ATGCCCAAGGCACTTCTGTGTGGGGCGCTGATCCACCCTGCTGAGCTGCTGCTGCTGCT 1080
QY 1081 GCCCATGACGACGACATTAACCAACCACTTAAGGTGGTGGAGGAGCCAGAGACCTG 1140
DB 1081 GCCCATGACGACGACATTAACCAACCACTTAAGGTGGTGGAGGAGCCAGAGACCTG 1140
QY 1141 AAGAAAGAAATTTTATGAGACAGACCTTTAGGGTGAAGATATTCAATPACAGCCAC 1200
DB 1141 AAGAAAGAAATTTTATGAGACAGACCTTTAGGGTGAAGATATTCAATPACAGCCAC 1200
QY 1201 TACATGAAGAGATGATGATTTCCCAATGATATTGATGCTCAAGTTAAAGACAGT 1260
DB 1201 TACATGAAGAGATGATGATTTCCCAATGATATTGATGCTCAAGTTAAAGACAGT 1260
QY 1261 GATGCTCATCTGCTTGAATCCAAATACGTGAAGACTGTGTCTTCCCTGATGGCTCC 1320
DB 1261 GATGCTCATCTGCTTGAATCCAAATACGTGAAGACTGTGTCTTCCCTGATGGCTCC 1320
QY 1321 TTTCCCTGAGAGTGAAGTGCACATCTGCTGGGGTGTTCACAGAAACAGAAAGAGG 1380
DB 1321 TTTCCCTGAGAGTGAAGTGCACATCTGCTGGGGTGTTCACAGAAACAGAAAGAGG 1380
QY 1381 TCCGCGACAGCTCCTGAGATGCCAAAGTCAAGCTGATTTGCCAACTTTGGCAACTCCGC 1440
DB 1381 TCCGCGACAGCTCCTGAGATGCCAAAGTCAAGCTGATTTGGCAACTTTGTGCACCTCCGC 1440
QY 1441 CAACTATAGCACCAATGATGATGACACTATGATCTGTGAGGAAATCTTCAGAAACT 1500
DB 1441 CAACTATAGCACCAATGATGATGACACTATGATCTGTGAGGAAATCTTCAGAAACT 1500
QY 1501 GGGGAGACACCTGACAGAGTGCCTGGAGGCCCTGACCTGTGAGAAAGCGGACAC 1560
DB 1501 GGGGAGACACCTGACAGAGTGCCTGGAGGCCCTGACCTGTGAGAAAGCGGACAC 1560
QY 1561 TACTACGCTATAGGATAGTACGCTGGGGCTGAGTGTGGAAAGAGCCAGGGCTTAC 1620
DB 1561 TACTACGCTATAGGATAGTACGCTGGGGCTGAGTGTGGAAAGAGCCAGGGCTTAC 1620
QY 1621 ACCCAAGTTACCAATTCCTGAATTTGATCAAGGACACATCAAAAGTGAAGTGGCTTC 1680
DB 1621 ACCCAAGTTACCAATTCCTGAATTTGATCAAGGACACATCAAAAGTGAAGTGGCTTC 1680
QY 1681 TAA 1683
DB 1681 TAA 1683

RESULT 2
ABN95170
ID ABN95170 standard; DNA: 3008 BP.
XX
AC ABN95170;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #1668 used to diagnose liver cancer.
XX
KM Gene: liver cancer; ds: hepatocellular carcinoma; hepatotropic;
KM metastatic liver tumour; cytostatic; expression profile: disease state;
KM disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.
XX
PN W0200229103-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US30589.
XX
PR 02-OCT-2000; 2000US-237054P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX
DR WPI: 2002-426119/45.
XX
PT Diagnosing and detecting the progression of liver cancer,
PT hepatocellular carcinoma and metastatic liver tumor in a patient.
PT Involves detecting the level of expression of two or more genes in a
PT liver tissue sample -
XX
XX
XX Claim 1: SEQ ID NO 1668; 298bp; English.
XX
XX The invention relates to a novel method for diagnosing and detecting the
XX progression of liver cancer, hepatocellular carcinoma or metastatic liver
XX tumour in a patient, and differentiating metastatic liver cancer from
XX hepatocellular carcinoma in a patient, involving detecting the level of
XX expression of two or more genes represented in ABN95170-ABN97455 in a
XX tissue sample. The method of the invention has hepatotropic, and
XX cytostatic activity. The method is useful for diagnosing and detecting
XX the progression of liver cancer, hepatocellular carcinoma and metastatic
XX liver carcinoma in a patient. The method is useful for identifying
XX expression profiles which serve as useful diagnostic markers as well as
XX markers that can be used to monitor disease states, disease progression,
XX drug toxicity, drug efficacy and drug metabolism.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 3008 BP: 791 A; 837 C; 703 G; 677 T; 0 other;
XX
XX Query Match 100.0%; Score 1683; DB 24; Length 3008;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 1683; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ATGTTGCCAGGATGTGATCTCATGTTCTGCTTAAATGCTGTGAGGAAGACA 60
DB 97 ATGTTGCCAGGATGTGATCTCATGTTCTGCTTAAATGCTGTGAGGAAGACA 156
OY 61 GCGTGGGGTCTGCTGATGCTTTATTTGGAAGCTGGAGCCAGACTGACCCCTGAC 120
DB 157 GCGTGGGGTCTGCTGATGCTTTATTTGGAAGCTGGAGCCAGACTGACCCCTGAC 216
OY 121 CAGATGATTTACAGCTACGAGGATTTATATCAGAGAAACACCACTAGACACTTACC 180
DB 217 CAGATGATTTACAGCTACGAGGATTTATATCAGAGAAACACCACTAGACACTTACC 276
OY 181 CATGCTGAATCTGACTGACTACTACACTGAGGAAAGCAAGCTGATCATCCAGCCCAAC 240

DB 277 CATGCTGAATCTGACTGACTACTACACTGAGGACCAAGCTGATCCATGCCAGCCCAAC 336
OY 241 CCCTGTGAACACGGGTGGGAGACTGCTGCTGTCATATGGGAGACACTTTCACATGACGCTGCTG 300
DB 337 CCCTGTGAACACGGGTGGGAGACTGCTGCTGTCATATGGGAGACACTTTCACATGACGCTGCTG 396
OY 301 GCTCTTCTCTCTGGGAATTAAGTGTGAGAAATGCAAAATATGCTGACAGACACCACTATG 360
DB 397 GCTCTTCTCTCTGGGAATTAAGTGTGAGAAATGCAAAATATGCTGACAGACACCACTATG 456
OY 361 GGCCTGGGCAATGCTCATTTACCCAGAGTCTCTCCATCTACCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 457 GGCCTGGGCAATGCTCATTTACCCAGAGTCTCTCCATCTACCTGCTGCTGCTGCTGCTGCTGCTG 516
OY 421 CCTTACACAGTCTCCAGCTGCTCCCAAGTGTCTCTGTATGACAGGCCAAACCCCTGCCAG 480
DB 517 CCTTACACAGTCTCCAGCTGCTCCCAAGTGTCTCTGTATGACAGGCCAAACCCCTGCCAG 576
OY 481 AATGGGCTACTGCTGCTCCGCGCATTAAGCCGAGATCCAAAGTTCACTGTGCTGCTGCTGCTGCTG 540
DB 577 AATGGGCTACTGCTGCTCCGCGCATTAAGCCGAGATCCAAAGTTCACTGTGCTGCTGCTGCTGCTG 636
OY 541 CAGTTCAGAGGAAATTTCTGGAATATGTTCTGATGATGCTGATGTTGCTGCTGCTGCTGCTGCTG 600
DB 637 CAGTTCAGAGGAAATTTCTGGAATATGTTCTGATGATGCTGATGTTGCTGCTGCTGCTGCTGCTG 696
OY 601 TCTTACCGAGGAAATTAAGTGTGAGAAATGCAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
DB 697 TCTTACCGAGGAAATTAAGTGTGAGAAATGCAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 756
OY 661 CACTTCTCTTTCAGAGAAATTTACACATGTTTATGAGAGTGTGTAACCCATGAGATTT 720
DB 757 CACTTCTCTTTCAGAGAAATTTACACATGTTTATGAGAGTGTGTAACCCATGAGATTT 816
OY 721 GGGGAACACATTTCTGAGAAACCCAGATGCGGACGAAAGCCCTGCTGCTTATTTAA 780
DB 817 GGGGAACACATTTCTGAGAAACCCAGATGCGGACGAAAGCCCTGCTGCTTATTTAA 876
OY 781 GTTACCAATGACAAAGTGAATGGAATGGAATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
DB 877 GTTACCAATGACAAAGTGAATGGAATGGAATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 936
OY 841 GTTGCCTACCCAGAGAAAGCCCACTGAGCCATCAACCAAGCTTCCGGGGTTTACATCC 900
DB 937 GTTGCCTACCCAGAGAAAGCCCACTGAGCCATCAACCAAGCTTCCGGGGTTTACATCC 996
OY 901 TGTGGAAGACTGAGATTAAGCAGAGAGAAATCAAGAGAATCTATGAGGCTTTAAAGAC 960
DB 997 TGTGGAAGACTGAGATTAAGCAGAGAGAAATCAAGAGAATCTATGAGGCTTTAAAGAC 1056
OY 961 ACGGCGGCAACACCCATGCGAGGCTGCTCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
DB 1057 ACGGCGGCAACACCCATGCGAGGCTGCTCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1116
OY 1021 ATGCCCCAGGCGCACTTGT 1080
DB 1117 ATGCCCCAGGCGCACTTGT 1176
OY 1081 GCCCATGCAACGACATTAATAACACAGACATCTTAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
DB 1177 GCCCATGCAACGACATTAATAACACAGACATCTTAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1236
OY 1141 AAGGAAGAAGATTTCAATGACAGCTTAAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
DB 1237 AAGGAAGAAGATTTCAATGACAGCTTAAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1296
OY 1201 TACAATGAAGAAGATGAGATTTCCCAATGATATTTGCAATGCTCAAGTTAAAGCCAGTG 1260
DB 1297 TACAATGAAGAAGATGAGATTTCCCAATGATATTTGCAATGCTCAAGTTAAAGCCAGTG 1356
OY 1261 GATGCTCACTGTGCTGTGAATCCAAATAGTGAAGACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320

```
Db 1357 GATGGTCACCTGCTCTAGAACCAATACATGAGAGACTGTCCTGATGGGTC 1416
QY 1321 TTTCCTCTGGAGTAGTGCACATCTCTGGCTGGGCTTTACAGAAACAGGAAAGG 1380
Db 1417 TTTCCTCTGGAGTAGTGCACATCTCTGGCTGGGCTTTACAGAAACAGGAAAGG 1476
QY 1381 TCCCGCCAGCTCTGGATGCCAAGCTGATGCGCAACACTTTGTGCAACTCCGCG 1440
Db 1477 TCCCGCCAGCTCTGGATGCCAAGCTGATGCGCAACACTTTGTGCAACTCCGCG 1536
QY 1441 CACGCTATGACCATGATGATGACAGTATGATCTGTGAGGAAATCTTCAGAACT 1500
Db 1537 CACGCTATGACCATGATGATGACAGTATGATCTGTGAGGAAATCTTCAGAACT 1596
QY 1501 GGGCAGACACCTGCGCAGGCTGACTGTGAGGCCCCCTGACCTGTGAGAGGCGCAC 1560
Db 1597 GGGCAGACACCTGCGCAGGCTGACTGTGAGGCCCCCTGACCTGTGAGAGGCGCAC 1656
QY 1561 TACTAGCTATGGAGATAGTACCTGGGCTGTGAGTGTGGAGAGGCCAGGGGCTTAC 1620
Db 1657 TACTAGCTATGGAGATAGTACCTGGGCTGTGAGTGTGGAGAGGCCAGGGGCTTAC 1716
QY 1621 ACCCAAGTTACCAATTCCTGATGATGATGATGATGATGATGATGATGATGATG 1680
Db 1717 ACCCAAGTTACCAATTCCTGATGATGATGATGATGATGATGATGATGATGATG 1776
QY 1681 TAA 1683
Db 1777 TAA 1779

RESULT 3
AAL45697
ID AAL45697 standard: DNA, 1683 BP.
AAL45697:
XX AC
XX 13-JUN-2002 (first entry)
XX DE Human blood coagulation factor VII activating protease mutant DNA.
XX KM Human; blood coagulation factor VII activating protease: FSAP;
XX KW single-chain plasminogen activator; bleeding disorder; haematological;
XX haemostatic; mutant; gene; ds.
XX OS Homo sapiens.
XX OS Synthetic.
XX FT Key Location/Qualifiers
XX FT CDS 1..1683
XX FT FT /*tag= a
XX FT FT /product= "mutant FSAP"
XX PN EPI182258-A1.
XX PD 27-FEB-2002.
XX PE 05-JUL-2001; 2001EP-0115691.
XX PR 26-JUL-2000; 2000DE-1036641.
XX PR 10-OCT-2000; 2000DE-1050040.
XX PR 21-OCT-2000; 2000DE-1052319.
XX PR 12-APR-2001; 2001DE-1018706.
XX PA (AVET ) AVENTIS BEHRING GMBH.
XX PI Roemisch J, Stoeck H, Feussner A, Lang W, Weimer T, Becker M;
XX PI Nerlich C, Muth-Naumann G;
XX DR WPI: 2002-270939/32.
XX DR P-PSDB: AAO17145.
XX PT New nucleic acid encoding mutant factor 7 activating protease, useful
```

for diagnosis, treatment and prevention of coagulation disorders, also related protein and antibodies

Disclosure: Page 16-17; 27p; German.

The present invention relates to a mutant of the DNA sequence encoding the protease (FSAP) that activates blood coagulation factor VII (FVII) and single-chain plasminogen activator, where at least one of the base changes G to C at nucleotide 1177 and G to A at nucleotide 1601 is present. The mutant sequences can be used in the treatment and prevention of bleeding disorders associated with inherited or acquired defects of blood coagulation factors V, VIII, IX, X, XI, XII, von Willebrand's factor and/or with antibodies against any of these proteins. The present sequence is the mutant human FSAP coding sequence.

Sequence 1683 BP; 441 A; 437 C; 435 G; 370 T; 0 other;

Query Match 99.8%; Score 1679.8; DB 24; Length 1683;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1681; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGTTGCCAGATGCTGATCTCCATGTTCTGTTAATGAGCTGTGGGAAAGACA 60
Db 1 ATGTTGCCAGATGCTGATCTCCATGTTCTGTTAATGAGCTGTGGGAAAGACA 60
QY 61 GCCTGTGGGTTCCCTGATGCTTTATTTGAAAGCTGAGCCAGACCTGCGAC 120
Db 61 GCCTGTGGGTTCCCTGATGCTTTATTTGAAAGCTGAGCCAGACCTGCGAC 120
QY 121 CAGTATGATTACAGCTACAGAGATTATATCAGAGAACACAGTACAGACTTACC 180
Db 121 CAGTATGATTACAGCTACAGAGATTATATCAGAGAACACAGTACAGACTTACC 180
QY 181 CATGCTAGATCTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 240
Db 181 CATGCTAGATCTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 240
QY 241 CCCGTGTAACACAGGTTGGGAGCTGCTGCTCATGAGACCTTTCATGATGAGCTG 300
Db 241 CCCGTGTAACACAGGTTGGGAGCTGCTGCTCATGAGACCTTTCATGATGAGCTG 300
QY 301 GCTCTTTCTCTGGGAATTAAGTGTGAGAAAGTCAAAATACGTGACAGACCACTGT 360
Db 301 GCTCTTTCTCTGGGAATTAAGTGTGAGAAAGTCAAAATACGTGACAGACCACTGT 360
QY 361 GGCCTGGGCTCAATGCTCATTAACCAAGTCTCTCTATGACAGCCCAACCCCTG 420
Db 361 GGCCTGGGCTCAATGCTCATTAACCAAGTCTCTCTATGACAGCCCAACCCCTG 420
QY 421 CCTTACACAGTCTCCAGCTGCTCCCAAGTGTCTGTATGACAGCCCAACCCCTG 480
Db 421 CCTTACACAGTCTCCAGCTGCTCCCAAGTGTCTGTATGACAGCCCAACCCCTG 480
QY 481 AATGGGGCTACCTGCTCCGCGATTAAGCGAGATTCACCTGCTGCTGCTGCTG 540
Db 481 AATGGGGCTACCTGCTCCGCGATTAAGCGAGATTCACCTGCTGCTGCTGCTG 540
QY 541 AATGGGGCTACCTGCTCCGCGATTAAGCGAGATTCACCTGCTGCTGCTGCTG 600
Db 541 AATGGGGCTACCTGCTCCGCGATTAAGCGAGATTCACCTGCTGCTGCTGCTG 600
QY 601 TCTTACCGAGGAAATGAAATAGACAGTCAACGCAATGCTGCTTACGGAATCC 660
Db 601 TCTTACCGAGGAAATGAAATAGACAGTCAACGCAATGCTGCTTACGGAATCC 660
QY 661 CACGTCCTCTGAGAGAAATTAACATGTTATAGAGATGCTGAACCCATGGGATT 720
Db 661 CACGTCCTCTGAGAGAAATTAACATGTTATAGAGATGCTGAACCCATGGGATT 720
QY 721 GGGGAACAAATTTCTGAGAAACCAAGTGGGAGGAAAGCCCTGCTTATTAATA 780
Db 721 GGGGAACAAATTTCTGAGAAACCAAGTGGGAGGAAAGCCCTGCTTATTAATA 780
```


[illegible]

XX Human ORFX ORF2248 polynucleotide sequence SEQ ID NO:4455.

KW Human: open reading frame; ORX: detection; cytoskeletal; hepatotropic;
KW vlntherapy; antipsoeopathic; antiparkinsonian; neurotropic; neuroprotective;
KW anticonvulsant; osteopathic; antilarthritic; immunosuppressant; cardiac;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antinaemic; gene therapy; cancer; proliferative disorder; hyperextension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antinflammatory disease; coagulation;
KW thrombosis; contraceptive; SS.

XX Homo sapiens.

XX XX

XX WO200058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US08621.

XX PF 31-MAR-1999; 99US-0127607.

XX PR 02-APR-1999; 99US-0127636.

XX PR 05-APR-1999; 99US-0127728.

XX PR 30-MAR-2000; 2000US-0540763.

XX (CURA-) CURAGEN CORP.

XX PA Shimkels RA, Leach M;

XX PI MPI: 2000-602362/57.

XX DR P-PSDB: AAB2484.

XX XX

XX Novel nucleic acids and peptides derived from open reading frame X,
XX useful for treating e.g. cancers, proliferative disorders,
XX neurodegenerative disorders and cardiovascular disease -
XX PT

XX Claim 5: Page 3679-3681; 5507pp; English.

XX XX

XX AAC74446 to AAC74606 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX sequences have activities such as: cytostatic; hepatotropic; vlntherapy;
XX antipsoeopathic; antiparkinsonian; neurotropic; neuroprotective;
XX osteopathic; anticonvulsant; antilarthritic; immunosuppressant;
XX immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
XX antidiabetic; hypotensive; dermatological; immunosuppressive;
XX antinflammatory; antibacterial; antirheumatic; antirheumatic;
XX antithyroid; and antinaemic. The sequences can be used for determining
XX the presence of or predisposition to, or preventing or treating
XX pathological conditions associated with an ORFX-associated disorder. The
XX nucleic acids can be used to express ORFX proteins in gene therapy,
XX vectors. The proteins and nucleic acids may be used to treat cancers,
XX proliferative disorders, neurodegenerative disorders, osteoarthritis,
XX graft vs host disease, cardiovascular disease, diabetes mellitus,
XX hyperextension, hypothyroidism, cholesterol ester storage, systemic lupus
XX erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
XX bacterial or fungal infection, malaria, autoimmune disorders, asthma,
XX allergic, aplastic anaemia, burns, wounds, bone and cartilage damage,
XX nocturnal haemoglobinuria, antinflammatory disease; to enhance
XX coagulation; to inhibit thrombosis; and as a contraceptive.

XX XX

XX Sequence 3623 BP; 848 A; 1005 C; 930 G; 840 T; 0 other;

XX Query Match 99.7%; Score 1678.2; DB 21; Length 3623;

XX Best Local Similarity 99.8%; Pred. No. 0;

XX Matches 1680; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

XX 1 ATCTTTCGAGGATGTCGATCCATCTTTCGCGCTTAATGCGCTGCTGGAGAACGA 60

Dh	97	ATGTTGGCAGGATGTGTCGATCTCCATGTTCTGCTGTAAAGGCTCTGTGTGGGAAGACA	156
Oy	61	GCTGTGGGTTCTCCCTGATGTTCTTATTGGAAGCTGGACCCAGACTGGACCTGTAC	120
Dh	157	GCGTGTGGGTTTCCCTCATGTCTTATTGGAAGGCTGGACCCAGACTGGACCTGTAC	216
Oy	121	CAGTATGATACAGCTACGAGAGATTATATCAGSAAAGAACCCAGTAGAGCACTTACC	180
Dh	217	CAGTATGATACAGCTACGAGAGATTATATCAGSAAAGAACCCAGTAGAGCACTTACC	276
Oy	181	CATGTGTGAATCTCGACTGGTACTACTGTGGACCAACACTGTATCTATGGCAGCCAAC	240
Dh	277	CACGTGTGAATCTCGACTGGTACTACTGTGGACCAACACTGTATCTATGGCAGCCAAC	336
Oy	361	GGCGGGGGCCAAATGTCTATTACCAAGACTCCTCCATCAGCCGTGTGTGTGAACAC	420
Dh	457	GGCGGGGGCCAAATGTCTATTACCAAGACTCCTCCATCAGCCGTGTGTGTGAACAC	516
Oy	421	CCCTACAGAGTCCCAAGCTGTCCCAAGTGTTCCTGTATGCAAGGCCAAACCCCTGCAG	480
Dh	517	CCCTACAGAGTCCCAAGCTGTCCCAAGTGTTCCTGTATGCAAGGCCAAACCCCTGCAG	576
Oy	481	AATGGGGGTACCTGTCCCGGCATTAAGCCGAGATCTCAAGTTCACTGTCCTGTCCGAC	540
Dh	577	AATGGGGGTACCTGTCCCGGCATTAAGCCGAGATCTCAAGTTCACTGTCCTGTCCGAC	636
Oy	541	CAGTTCAAGGGGAAATCTCGAATAATAGTTCGTGTATGCTCATGTGTGGAGGTCATC	600
Dh	637	CAGTTCAAGGGGAAATCTCGAATAATAGTTCGTGTATGCTCATGTGTGGAGGTCATC	696
Oy	601	TCTTACGAGGGGAAAATGATATAGACAGTCAACAGCATGCGTGTCTTACTGGAACATCC	660
Dh	697	TCTTACGAGGGGAAAATGATATAGACAGTCAACAGCATGCGTGTCTTACTGGAACATCC	756
Oy	661	CACCTCTCTTGGAGAGATTAACAATGTTTATGAGAGATCTGAAACCACTAGGATT	720
Dh	757	CACCTCTCTTGGAGAGATTAACAATGTTTATGAGAGATCTGAAACCACTAGGATT	816
Oy	721	GGGGAACCAATTTCTGACGAACCCACAGATGGCGACGAAAGCCTGGTCTTATTAA	780
Dh	817	GGGGAACCAATTTCTGACGAACCCACAGATGGCGACGAAAGCCTGGTCTTATTAA	876
Oy	781	GATTACCATGACAAAGTGTGAATATGTAATCTGTATGTCTCAGCCGTCTAGGCCAGGAC	840
Dh	877	GATTACCATGACAAAGTGTGAATATGTAATCTGTATGTCTCAGCCGTCTAGGCCAGGAC	936
Oy	841	GTTGGCTTCCCAAGGAAGAGCCCACTGAGGCCATACCAAGTTCCGGGGTTTGAATCC	900
Dh	937	GTTGGCTTCCCAAGGAAGAGCCCACTGAGGCCATACCAAGTTCCGGGGTTTGAATCC	996
Oy	901	TGTGAAAGACTGATAGACAGAGAGAGATCAAGAGATCTATGAGGCTTAAAGAC	960
Dh	997	TGTGAAAGACTGATAGACAGAGAGAGATCAAGAGATCTATGAGGCTTAAAGAC	1056
Oy	961	ACGGGGGGCAGACCACTAGAGAGGCTCCCTCCAGTCTGTGCTGTATACCATCTCC	1020
Dh	1057	ACGGGGGGCAGACCACTAGAGAGGCTCCCTCCAGTCTGTGCTGTATACCATCTCC	1116
Oy	1021	ATGCCCCAGGGCCACTTGTGTGTGTGGGGCGTGATACCAACCCGTGTGTGTCTACGTCT	1080
Dh	1117	ATGCCCCAGGGCCACTTGTGTGTGTGGGGCGTGATACCAACCCGTGTGTGTCTACGTCT	1176
Oy	1081	GGCCACTGCACCGACATTAATAAACCAAGCATCTTAAAGGTGTGTGTAGGGACACGAGCTTG	1140
Dh	1177	GGCCACTGCACCGACATTAATAAACCAAGCATCTTAAAGGTGTGTGTAGGGACACGAGCTTG	1236

QY	1141	AAGAAACAAGATTTCATGACGACGAGATTTCAGGTGGTGGACAGATATTTCAATTCAGTACAGCCAC	1200
Db	1237	AAGAAACAAGATTTCATGACGACGACCTTTAGGGTGGAGAGATATTTCAAGTACAGCCAC	1296
QY	1201	TACAAATGAAGAAGATGACATTTCCCAACATGATATTTGCATTGCTCTCAAGTTAAAGCCAGTG	1260
Db	1297	TACAAATGAAGAAGATGACATTTCCCAACATGATATTTGCATTGCTCTCAAGTTAAAGCCAGTG	1356
QY	1261	GATGGTCACTGTGCTCTGATGAAATCCAAATACCTGGAAGACGTGTGCTTGCCTGATGGATCC	1320
Db	1357	GATGGTCACTGTGCTCTGATGAAATCCAAATACCTGGAAGACGTGTGCTTGCCTGATGGATCC	1416
QY	1321	TTTCCCTCTGGGGAGTGGATGGCCATCTCTGTGGTGGGGTGGTTACAGAAACAGCAAAACGG	1380
Db	1417	TTTCCCTCTGGGGAGTGGATGGCCATCTCTGTGGTGGGGTGGTTACAGAAACAGCAAAACGG	1476
QY	1381	TCCCGCCAGCTCCCTGGATGCCAAAGTCAGCTGATTGGCCACACCTTTGTGCATCTCCGC	1440
Db	1477	TCCCGCCAGCTCCCTGGATGCCAAAGTCAGCTGATTGGCCACACCTTTGTGCATCTCCGC	1536
QY	1441	CACCTCTATGACCACATGATTGATGTGACAGTATGATCTGTGTGACAGAAATCTTCAGAAACCT	1500
Db	1537	CACCTCTATGACCACATGATTGATGTGACAGTATGATCTGTGTGACAGAAATCTTCAGAAACCT	1596
QY	1501	GGGCAAGACACCGGACGAGGTCGTGTGGAGGCCCCCTGACCTGTGGAAGAGCGACGC	1560
Db	1597	GGGCAAGACACCGGACGAGGTCGTGTGGAGGCCCCCTGACCTGTGGAAGAGCGACGC	1656
QY	1561	TACTTACTCTATGGGATATGTAGTACTGGGGCTGGAGTGTGGGAGAGAGCCAGGGGTCTAC	1620
Db	1657	TACTTACTCTATGGGATATGTAGTACTGGGGCTGGAGTGTGGGAGAGAGCCAGGGGTCTAC	1716
QY	1621	ACCCAAAGTACCAAAATTCCTGAATTGGATCAAAAGCCACCATCAAAAGTGAAGATGGCTTC	1680
Db	1717	ACCCAAAGTACCAAAATTCCTGAATTGGATCAAAAGCCACCATCAAAAGTGAAGATGGCTTC	1776
QY	1681	TAA 1683	
Db	1777	TAA 1779	
RESULT 5			
AAAA4763			
ID	AAAA4763 standard; cDNA; 617 BP.		
XX	AAAA4763;		
XX	21-AUG-2000 (first entry)		
DE	Human secreted expressed sequence tag SEQ ID NO:1338.		
XX			
KW	Human; mouse; chicken; rat; secreted expressed sequence tag; SESTF;		
KW	expressed sequence tag; EST; probe; chemotactic; proliferative;		
KW	immunomodulatory; hematopoietic; chemokinetic; analgesic; pneumostatic		
KW	thrombolytic; antiinflammatory; cytosolic; antibacterial; antifungal;		
KW	antiviral; antidiabetic; antiproteolytic; vulnerary; antiparkinsonian;		
KW	antibacter; osteopathic; neuroprotective; nootropic; antiproliferative;		
KW	cardioprotective; anticonvulsant; antidepressant; gene therapy;		
KW	vaccines; autoimmune disorder; multiple sclerosis; allergic condition;		
KW	insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;		
KW	lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;		
KW	central nervous system disorder; Alzheimer's disease; stroke;		
KW	Parkinson's disease; Huntington's disease; coagulation disorder;		
KW	hemophilia; thrombosis; inflammatory disorder; Crohn's disease;		
KW	tumour; infection; depression; psoriasis; ss.		
OS	Homo sapiens.		
XX			
XX	MO2000021991-A1.		
XX			
XX	20-APR-2000.		
XX			

```
PF 15-OCT-1999; 99MO-US24206.
XX
PR 15-OCT-1998; 98US-0104436.
XX
PA (GENM ) GENETICS INST INC.
XX
PI Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;
PI Merberg D, Treacy M, Bowman MR;
XX WPI: 2000-317938/27.
XX
PT Isolated polynucleotides, and encoded proteins, comprising secreted
PT expressed sequence tags (SESTs), useful for treating various disorders
PT such as autoimmune, infectious, and central nervous system disorders -
XX
XX Claim 1: Page 549; 803pp: English.
XX
CC AAA43426 to AAA45925 represent specifically claimed secreted expressed
CC sequence tags (SESTs), isolated from human, mouse, chicken and rat
CC tissue sources. The SESTs can have a range of activities depending on
CC the tissues they were isolated from. The activities include:
CC chemokinetic; proliferative; immunomodulatory; haematopoietic;
CC cytostatic; antibacterial; antifungal; antiviral; antidiabetic;
CC antileukemic; antitumor; antitumor; osteoprotective; neuroprotective;
CC anticonvulsant; and antidepressant. The SESTs can be used for gene
CC therapy and in vaccines. The SESTs are useful as probes for the
CC identification and isolation of full-length cDNAs and genomic DNA
CC molecules which correspond to the SESTs. Proteins encoded by the SESTs
CC are useful in assays for determining biological activity and raising
CC antibodies. They may be useful for treatment of autoimmune disorders
CC (multiple sclerosis, insulin dependent diabetes), allergic conditions
CC (asthma), myeloid or lymphoid cell deficiencies, wounds, ulcers,
CC osteoporosis, osteoarthritis, central nervous system disorders,
CC Alzheimer's, Parkinson's, Huntington's disease, stroke, coagulation
CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's
CC disease), tumours, bacterial, fungal or viral infections, depression and
CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given
CC in the exemplification of the present invention.
XX
SQ Sequence 617 BP; 165 A; 169 C; 148 G; 135 T; 0 other;
Query Match 31.3%; Score 526.4; DB 21; Length 617;
Best Local Similarity 99.8%; Pred. No. 2.8e-142;
Matches 527; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
DB 444 GGCGGGGCAATGTCATTACCCAGAGTCTCCCTACTACCTACCGCTGTGTAAACAC 503
QY 421 CCTACACAGGTCCACCTCCCAAGGTCTCTGTATGACGGCAACCCGCGAG 480
DB 504 CCTACACAGGTCCACCTCCCAAGGTCTCTGTATGACGGCAACCCGCGAG 563
QY 481 AATGGGGCTACCTCCCGGCAATTAAGCGAGATCCAAATTCACCTGT 528
DB 564 AATGGGGCTACCTCCCGGCAATTAAGCGAGATCCAAATTCACCTGT 611

RESULT 6
AB199281
ID AB199281 standard; cDNA; 397 BP.
XX
AC AB199281;
XX
DT 07-MAR-2002 (first entry)
XX
DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:123.
XX
KW Mouse: ischaemia; compressive ischaemia; occlusive ischaemia;
KW vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.
XX
OS Mus musculus.
XX
PN W020018188-A2.
XX
PD 22-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-JP04192.
XX
PR 18-MAY-2000; 2000JP-0145977.
XX
PA (UYN1-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX WPI: 2002-034733/04.
XX
DR
XX
PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or
PT by determining the expression profile of a gene group comprising these
PT genes -
XX
PS Claim 2: Page 351-352; 2690pp: English.
XX
XX The present invention describes a method for examining ischemic
XX conditions, comprising measuring the expression levels of particular
XX genes (I) in a test sample or determining the expression profile of a
XX gene group in the sample comprising genes selected from (I). The method
XX is useful for examining the ischemic condition (e.g. compressive
XX ischemia, occlusive ischemia or vasospastic ischemia) by measuring
XX expression levels of particular genes (AB199202 to AB199912, encoding
XX the protein sequences in AB857020 to AB857374) or by determining the
XX expression profile of a gene group comprising these genes. The
XX expression levels or expression profiles produced by these genes are
XX used as an indicator when screening for ischemic condition-improving
XX drugs or as a predictor for ischemic diseases. AB19913 and AB19914
XX represent PCR primers for a mouse ischemic condition related sequence,
XX which are used in the exemplification of the present invention.
XX
SQ Sequence 397 BP; 103 A; 105 C; 111 G; 78 T; 0 other;
Query Match 14.4%; Score 242.6; DB 24; Length 397;
Best Local Similarity 84.7%; Pred. No. 5.2e-60;
Matches 272; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
```

```
|||||
Db 73 CCTTGTGCAACTCCGACACACTATGACCAACCATCATGATGATTTGTGCG 132
1483 GGAATCTCTGAAACACCTGGGACACCTGGAGGTGACTCTGAGGCCCTTGACC 1542
133 GGGAACTCTGAAAGCCGGATCAGACACTGCGGAGGAGGAGGAGGAGGAGGAGG 192
1543 TGTGAGAGGACGACACTACTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 1602
193 TGTGAGAGGATGAGACTACTACTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAG 252
Qy 1603 AAGAGCCGAGGAGTCTACACCAAGTTCGAAATTCGAAATTCGAAATTCGAA 1662
253 AAGAGCCGAGGAGTCTACACCAAGTTCGAAATTCGAAATTCGAAATTCGAAATTC 312
Db 1663 AAAAGTGAAGTGGCTTCTAA 1683
313 CACAGGAGGAGTGGCTTCTGA 333

RESULT 7
ABAS8323/c
ID ABAS8323 standard; DNA; 451 BP.
XX
AC ABA58323;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #6628.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
PD 09-AUG-2001.
XX
PE 30-JAN-2001; 2001WO-US00669.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0632366.
XX
PR 21-SEP-2000; 2000US-0234687.
XX
PR 27-SEP-2000; 2000US-0236359.
XX
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for
XX
PT analyzing gene expression in human fetal liver -
XX
PS Claim 1; SEQ ID NO 6628; 639pp + sequence listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC foetal liver. The present sequence is a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pcr_sequences.
XX
SQ Sequence 451 BP; 116 A; 102 C; 127 G; 106 T; 0 other;

Query Match 7.1%; Score 120.2; DB 22; Length 451;
Best Local Similarity 97.6%; Pred. No. 1.8e-24;
```

```
Matches 122; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 324 TCAGAAAGTCAAAATACGTGACAGACCAACCATGTGGCGGGGCAATGTCTCATTC 383
451 TCCTACAGTCCGAAATACGTGACAGACCAACCATGTGGCGGGGCAATGTCTCATTC 392
Qy 384 CCAGAGTCCCTCTACTACGCGGTGTGTGTAAACACCCCTTACACAGTCCAGTGTTC 443
391 CCAGAGTCCCTCTACTACGCGGTGTGTGTAAACACCCCTTACACAGTCCAGTGTTC 332
Db 444 CCAAG 448
331 CCAAG 327

RESULT 8
AAK06420/c
ID AAK06420 standard; DNA; 451 BP.
XX
AC AAK06420;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 6411.
XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
PD 09-AUG-2001.
XX
PE 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0632366.
XX
PR 21-SEP-2000; 2000US-0234687.
XX
PR 27-SEP-2000; 2000US-0236359.
XX
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
XX
PT brains -
XX
PS Example 4; SEQ ID NO: 6411; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention.
XX
SQ Sequence 451 BP; 116 A; 102 C; 127 G; 106 T; 0 other;

Query Match 7.1%; Score 120.2; DB 22; Length 451;
Best Local Similarity 97.6%; Pred. No. 1.8e-24;
Matches 122; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 324 TCAGAAAGTCAAAATACGTGACAGACCAACCATGTGGCGGGGCAATGTCTCATTC 383
451 TCCTACAGTCCGAAATACGTGACAGACCAACCATGTGGCGGGGCAATGTCTCATTC 392
```

QY 384 CCAGAGTCCTCCCTACTACCGCTGTGTCTGTAACACCCCTTACAGAGTCCAGGCTGTC 443
 Db 391 CCAGAGTCCTCCCTACTACCGCTGTGTCTGTAACACCCCTTACAGAGTCCAGGCTGTC 332

QY 444 CCAAG 448
 Db 331 CCAAG 327

RESULT 9

AAK32096/C
 ID AAK32096 standard; DNA: 451 BP.

AC AAK32096;

DT 06-NOV-2001 (first entry)

DE Human bone marrow expressed single exon probe SEQ ID NO: 6653.

KM Human: bone marrow expressed exon: gene expression analysis: probe:

KM microarray: cancer; leukemia; lymphoma; myeloma; ss.

OS Homo sapiens.

PN WO200157276-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00668.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR:

XX WPI: 2001-488900/53.

PT Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human bone marrow -

PS Example 4: SEQ ID NO: 6653; 658bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC bone marrow. They can be used to measure gene expression in bone marrow

CC samples, which may enable the improved diagnosis and treatment of cancers

CC such as lymphoma, leukemia and myeloma. The present sequence is one of

CC the probes of the invention.

XX Sequence 451 BP; 116 A; 102 C; 127 G; 106 T; 0 other:

QY 324 TCAGAAAGTCGCAAGAGCAACCAATGTCGCGGGCCATGTCATTAC 383

Db 451 TCCTACAGTCGCAAAATACGTGCAAGAGCAACCAATGTCGCGGGCCATGTCATTAC 392

QY 384 CCAGAGTCCTCCCTACTACCGCTGTGTCTGTAACACCCCTTACAGAGTCCAGGCTGTC 443

Db 391 CCAGAGTCCTCCCTACTACCGCTGTGTCTGTAACACCCCTTACAGAGTCCAGGCTGTC 332

QY 444 CCAAG 448

Db 331 CCAAG 327

RESULT 10
 ID AAI37947/C
 AAI37947 standard; DNA: 451 BP.

AC AAI37947;

DT 17-OCT-2001 (first entry)

DE Probe #6633 used to measure gene expression in human placenta sample.

KM Probe: microarray; human; placenta; antenatal diagnosis;

KM genetic disorder; ss.

OS Homo sapiens.

PN WO200157272-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00663.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR:

XX WPI: 2001-48897/53.

PT Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human placenta -

PS Claim 25; SEQ ID NO 6633; 654bp; English.

XX The present invention relates to single exon nucleic acid probes (SENP).

CC The present sequence is one such probe. The probes are useful for

CC producing a microarray for predicting, measuring and displaying gene

CC expression in samples derived from human placenta. The probes are useful

CC for antenatal diagnosis of human genetic disorders.

XX Sequence 451 BP; 116 A; 102 C; 127 G; 106 T; 0 other:

QY 324 TCAGAAAGTCGCAAGAGCAACCAATGTCGCGGGCCATGTCATTAC 383

Db 451 TCCTACAGTCGCAAAATACGTGCAAGAGCAACCAATGTCGCGGGCCATGTCATTAC 392

QY 384 CCAGAGTCCTCCCTACTACCGCTGTGTCTGTAACACCCCTTACAGAGTCCAGGCTGTC 443

Db 391 CCAGAGTCCTCCCTACTACCGCTGTGTCTGTAACACCCCTTACAGAGTCCAGGCTGTC 332

QY 444 CCAAG 448

Db 331 CCAAG 327

RESULT 11

ABS06863/C
 ID ABS06863 standard; DNA: 451 BP.

XX ABS06863;

XX 19-AUG-2002 (first entry)
 DT Human genome-derived single exon probe from lung SEQ ID NO 6854.
 XX
 XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary hemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Kargener syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 KM
 KM
 OS Homo sapiens.
 XX
 XX
 XX WO200186003-A2.
 PM
 PM 15-NOV-2001.
 PD
 PD 30-JAN-2001; 2001WO-US00665.
 PF
 PF 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608409.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI
 DR WPI: 2002-114183/15.
 XX
 XX Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 PS
 PS Claim 1: SEQ ID NO 6854; 634bp; English.
 XX
 XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray, assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived RNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary

CC	haemodieresis, pulmonary histiocytosis, lymphangioleiomyomatosis,
CC	pulmonary alveolar proteinosis, Kargener syndrome, fibrocystic
CC	pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
CC	and hyaline membrane disease. The present sequence is a single exon
CC	probe of the invention.
CC	Note: The sequence data for this patent did not form part
CC	of the printed specification, but was obtained in electronic
CC	format directly from WIP0 at
CC	ftp.wipo.int/pub/published_pcl_sequences.
XX	
XX	Sequence 451 BP: 116 A; 102 C; 127 G; 106 T; 0 other:
SQ	
Query Match	7.1%; Score 120.2; DB 24; Length 451;
Best Local Similarity	97.6%; Pred. No. 1.8e-24;
Matches 122:	Conservative 0; Mismatches 3; Indels 0; Gaps
OY	324 TCAGAAAGCGCAAAATACGCGAAGGACACCCATGTGGCGGGGCAATGTCATTTAC 383
DB	451 TCCTACAGCGCAAAATACGCGAAGGACACCCATGTGGCGGGGCAATGTCATTTAC 392
OY	384 CCAGAGTCTCTCCCTACTACCGCGTGTCGTGTAACACCCCTTAACAGGTCGACGTCTC 443
DB	391 CCAGAGTCTCTCTCTACTACCGCGTGTCGTGTAACACCCCTTAACAGGTCGACGTCTC 392
OY	444 CCAAG 448
DB	331 CCAAG 327
RESULT 12	
AA063951	ID
AA063951	standard; cDNA; 2033 BP.
XX	
AC	AA063951;
XX	
DT	06-JAN-1995 (first entry)
XX	
DE	Hepatocyte growth factor converting protease coding sequence.
XX	
KW	Hepatocyte growth factor; protease; cleavage; active; inactive;
KW	precursor; ds.
XX	
OS	Homo sapiens.
XX	
PH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..1968
FT	/tag= a
FT	/product= Hepatocyte growth factor converting
FT	protease.
XX	
PN	EP596524-A.
XX	
PD	11-MAY-1994.
XX	
PP	05-NOV-1993; 93BP-0117988.
XX	
PR	05-NOV-1992; 92JP-0296133.
PR	20-NOV-1992; 92JP-0312234.
PR	20-NOV-1992; 92JP-0312242.
XX	
PA	(SHIM/) SHIMOMURA T.
PA	(MITU) MITSUBISHI KASEI CORP.
XX	
PI	Kitamura N, Miyazawa K, Morimoto Y, Shimomura T;
XX	Yamada K;
XX	
XX	WPI; 1994-152921/19.
XX	DR P-PSDB; AAR53962.
PT	Hepatocyte growth factor converting protease and precursor and
PT	gene encoding them - for producing active two chain HGF from
PT	inactive single chain HGF
XX	

Query Match	7.1%;	Score 120.2;	DB 24;	Length 451;
Best Local Similarity	97.6%;	Pred. No. 1.0e-24;		
Matches 122;	Conservative	0;	Mismatches 3;	Indels 0;
				Gaps 0;
OY	324	TCGAGAAAGTCGAAATACGTGCAGAGACACCCATGTGGCGCGGGGCCAATGTCATTCAC	383	
Db	451	TCCTACAGTCGAAATACGTGCAGAGACACCCATGTGGCGGGGCCAATGTCATTCAC	392	
OY	384	CCAGAGTCCTCCCTACAGCGCGTGTCGTAAACACCCCTTACACAGGTCACAGTCTC	443	
Db	391	CCAGAGTCCTCCCTACAGCGCGTGTCGTAAACACCCCTTACACAGGTCACAGTCTC	332	
OY	444	CCAGA 448		
Db	331	CCAGA 327		
RESULT 12				
AA063951	ID	AA063951 standard; CDNA; 2033 BP.		
XX	AC	AA063951;		
XX	DF	06-JAN-1995 (first entry)		
XX	DE	Hepatocyte growth factor converting protease coding sequence.		
XX	KW	Hepatocyte growth factor; protease; cleavage; active; inactive;		
XX	KN	precursor; ds.		
XX	OS	Homo sapiens.		
XX	EH	Key	Location/Qualifiers	
XX	FT	CDS	1..1968	
XX	FT		/*tag= a	
XX	FT		/product= Hepatocyte growth factor converting	
XX	FT		protease.	
XX	PN	EP596524-A.		
XX	PD	11-MAY-1994.		
XX	PF	05-NOV-1993;	93BP-0117988.	
XX	PR	05-NOV-1992;	92JP-0296133.	
XX	PR	20-NOV-1992;	92JP-0312234.	
XX	PR	20-NOV-1992;	92JP-0312242.	
XX	PA	(SHIM/) SHIMOMURA T.		
XX	PA	(MITU) MITSUBISHI KASEI CORP.		
XX	PI	Kitamura N, Miyazawa K, Morimoto Y, Shimomura T;		
XX	PI	Yamada K;		
XX	DR	WPI; 1994-152921/19.		
XX	DR	P-PSDB; AAR53962.		
XX	PT	Hepatocyte growth factor converting protease and precursor and		
XX	PT	gene encoding them - for producing active two chain HGF from		
XX	PT	inactive single chain HGF		
XX				

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp://wipo.int/pub/published_pct_sequences.

XX Sequence 2036 BP; 339 A; 743 C; 623 G; 331 T; 0 other;

Query Match 7.0%; Score 117.6; DB 24; Length 2036;
Best Local Similarity 49.8%; Pred. No. 2.2e-23;
Matches 607; Conservative 0; Mismatches 359; Indels 54; Gaps 10;

```

469 AACCCCTGCGAATGGGGTACTGCTCCGGCATAAGCGAGATCCAGTTACCTGT 528
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 745 ACCCTTCTCGAAGGGGGACCTCCACCTGATCGGGCCACGGAGACCGTGT 804

529 GCCTGCCGACGAGTTCAAGGGAAATCTGTGAATAGTTGTGTA---CTGCTAT 585
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 805 GCCTGCCGACGAGCTTCTGAGCGGCTGTCAACATCGAGCTGATGAGCGCTTTC 864

586 GTTGGCGATGGCTACTCTTACCGAGGAAATGAATAGACATCAACAGCATGCTGC 645
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 865 TTGGGGAAACGGCATGGGTACTGGCGTGGCCAGCACCCTCGGGCTCAAGCTGC 924

646 CTTTACTGGAACCTCCACCTCCCTTCCAGGAGATTAACACATGTTATGAGAGTGT 705
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 925 CTGGGCTGGAACCTCGATGCTGTCTACCGAGAGCTGCACGTGGACTCCGTGGCGCCG 984

706 GAACCCATGGGATTTGGGGAACAAATTTCTGAGAAAACCCAGATGGGAGAAAGCCC 765
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 985 GCCCTGTGGGGCTGGGGCCCCCATGCTTACTGCCGAATCCGGAATGACGAGAGAGCCC 1044

766 TGGTCTTTAATAAGTTACCATGACAAAGTGAATGGGAATCTGATGTCATGACC 825
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1045 TGGTCTGAGT---GGTGAAGAGAGCGGCTCTCTCTGGAGTACTGCCCTCGAGGCC 1101

826 TGCTGAGCCAGAGAGTTGGCTACCCAGAGAGAAAGCCCACTGAGCCATCAACCAAGCTT 885
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1102 TGGGAATCCCT-----CACGAGATCCAACTGTACCGGATCTCTGGCGACCTCG 1152

886 CCGGGCTTTGACTCCTGTGGAAGAGCTGAGATAGCAGAGAGAGATCAAGACATCTAT 945
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1153 CCTGAGCCAGCTCTCCCGGGGCGCAGGCTGCGGAGAGAGGACAAAGAGAGAGCTTC 1212

946 GGAAGCTTTAAGAGACGCGCGGCAAGACCCATGAGGAGCGCTCCCTGAGTCTGCTG 1005
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1213 CTGCGGCCACGATGATGAGGCGGCTCTCTCTGCTGCGGCTGCGACACCTG----- 1265

1006 CCTGTCACCATCTCCAGCCCGGACACTTCTGTGGTGGGGCGTATGATCCACCCCTGC 1065
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1266 --GCTGGCGGCATTTACATGCGGAGAGTTCTGCGCGGAGGCTGTGTCACACCTGC 1323

1066 TGGGTGCTACTGCTGCCACTGCAACGACATTAATAAACCAACATCTTA-----AGGTG 1119
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1324 TGGGTGCTGCGCGCGCCACTGCTTCTCCACAGCCGCCCGAGGAGAGGCTGCGTG 1383

1120 GTGCTAGGGGACCGAGACCTGAAGAAGAATTTCAATGACAGAGGTTTAGGGTGAAG 1179
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1384 GTGCTAGGGGACCGAGACCTTTTAACCGAGAGCGAGAGCTGACCGACCTTGGAGTGAAG 1443

1180 AAGATATCAAGTACGACGACCTACATGAANAAGATGAGATTCCCAACATGATATTGCA 1239
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1444 AAGTATATCCCGTACACCTGTAAGTGTTCACCCAGGAGCAC--GACCTGTGTC 1500

1240 TTGCTCAATTAAGCAAGTGAATGATGATGATGCTGCTGATGATCAATTAAGTCAAGACT 1299
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1501 CTGATCCGCGCTAATAAAGAAAGGGGACCGCTGTGCGACACGCTGCAGATTTGTGCAAGCC 1560

1300 GTGTGCTTGCCTGAT-----GGGTCTTTTCCCTGAGGAGTGAAGTGCACATCTGTGC 1353
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1561 ATCTGCTCGCCAGACCGGACGACACTTCCCGCAGGACACAACTGACCAATTTGCGGGG 1620

1354 TGGGG-----TGTTACAGAAACAGGAAAGGTTCCCGGACACTCTCTGTGATGCCAAAGTC 1407
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1621 TGGGGCACTTGTGATGAGAACGTGACGGGCTACTCCAGCTCCCTGCGGGAGGCCCTGTGTC 1680
```

```

Oy 1408 AACCTGATGCCAACACTTGTGCAACTCCCGCACTCTATGACCAATGATTGATGAC 1467
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1681 CCCCTGTGCGCCGACACACAAAGTGCAGAGCTGTAGAGTCTAGGGGCGGACATCAGCCCC 1740

1468 AGTATGATCTGTGACGAAATCTTCAGAAACCTGGCAAGACACCTGCCAGGGTGTACT 1527
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1741 AACATCTGTGTGCGGGCTACTT---GCACTGCATGTCGACGCGCTGCAGGGGAGCTCA 1797

1528 GAGAGCCCCCTTACTCTGTGAGAAGAGAGGACCTTACTAGTGTATGAGATAGAGAGCTGG 1587
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1798 GGGGGGCCCCCTGCGCTGCGAAGAGAACGCGGTGCTTACTCTTACGGCATATCAGCTGG 1857

1588 GCGCTGAGCTGAGG-----AGAGGCGAGGGGTCTACACCCCAAGTTACCAATTTCTTG 1641
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1858 GTGAGAGCGCTGGGGGCGGCTCCACAAAGCCGGGGGTCTACACCCGCGTGGCAACTATGTG 1917

Oy 1642 AATTGATCAAGCCACCAT 1661
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1918 GACTGATCAACGACCGGAT 1937

RESULT 14
ABL65242
ID ABL65242 standard; DNA; 2036 BP.
XX
AC ABL65242;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:3579.
XX
KW Human; cancer; colon; breast; ovary; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
```


PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237284P.
 PR 02-OCT-2000; 2000US-237285P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 XX
 PA (AVAL-) AVALON PHARM.
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Sopel DR, Weaver Z;
 XX
 XX WPI: 2002-188264/24.
 DR
 XX Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -
 XX
 PS Claim 1: SEQ ID 3579; 44pp; English.
 XX
 XX The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL6164
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.
 CC
 XX
 XX Sequence 2036 BP; 339 A; 743 C; 623 G; 331 T; 0 other:
 SQ
 Query Match 7.0%; Score 117.6; DB 24; Length 2036;
 Best Local Similarity 49.8%; Pred. No. 2.2e-23;
 Matches 607; Conservative 0; Mismatches 559; Indels 54; Gaps 10;
 QY 469 AACCCCTCCAGATGGGCTACCTGCTCCGCGATAGCGGAGATCCAACTTCACTGT 528
 DB 745 AACCTTGCCTGAAAGGGGGGACCTGCCACCTGATCTGGCCACCGGACCACTGCTGT 804
 QY 529 GCGTGTCCCGACGATCAAGGGGAAATTCGTGAATAGTTCGTGATGA---CTGCAT 585
 DB 805 GCGTGTCCCGACGATCTTCTGAGCGGCTCTGCAATCTCAGCCTTATGAGCGCTTC 864
 QY 586 GTTGGCGATGCTACTTCTACGAGGAAATGAAATGAGACATCAACGATGCTGC 645
 DB 865 TTGGGGAACGGACATGATGCTGAGCGTGGCCAGCACTCCGCGCTTCAGCTGC 924
 QY 646 CTTTATGGAACCTCCACCTCTCTTTCAGAGAGAAATTAACATGTTTATGAGAGATGCT 705
 DB 925 CTGGCTGGAACCTCGATCTGCTTACCAAGAGCTGACATGACCTCCCTGGCGCGCG 984
 QY 706 GAAACCCATGSGATTTGGGAGACACATTTCTGAGAAACCAATCGGACGAAAGCC 765
 DB 985 GCGCTGCTGGGCTGGGCGCCCTGCTACTGCGGATCCGGAATCCGACAGAGGCGC 1044
 QY 766 TGGTCTTTATTAAGTACCAATGACAAGGTGAATGGAATATCTGATGCTCAACC 825

DB 1045 TCGTCTACGT---GGTGAAGGACAGCGCGCTTCTCGGAGATGATGGCGCTGGAGGCC 1101
 QY 826 TCGTCTACGTGAGCGATGCTTCTTACCAAGAGAAAGCCCACTGAGCATCAACCACTT 885
 DB 1102 TCGGAATCCCT-----CACAGAGTCCAACTGTCACTGGATCTCTGGCCACCTTG 1152
 QY 886 CCGGGGTTTGAATCTGTGTGAAAGACTGAGATAGCAGAGAGAAATCAAGAGATCTAT 945
 DB 1153 CTTGAGCAGGCTTCCCGGGGGGCGAGGCTGGCGGAGGAGACAAAGAGAGACCTTC 1212
 QY 946 GAGGCTTTAAAGACACGCGGAGGAGCAAGCAGCAGGAGGCGCTCCACTGCTCCGTC 1005
 DB 1213 CTGGGCGCACGTTATCATCGCGGCTCTCTCTGCTGCGCGCTGACACCTCG----- 1265
 QY 1006 CCTCTGACCATCTTCATGCCCCCAAGGCCCATCTTGTGTGGGGGCGCTGATCCACCCTGC 1065
 DB 1266 --GCTGGCGCGCATCTACATCGGGAGACAGCTTGTGCGCGCGGAGCTGTCCACACCTGC 1323
 QY 1066 TGGGTGCTCACTGTGCGCCACATGACACGACATTAATAAACAGACATCTAA-----AGGTG 1119
 DB 1324 TGGGTGCTGTGGCGCGCCACATGCTTTCACAGCCCCCAGGAGACAGGCTCCGCTG 1383
 QY 1120 GTGCTAGGGGACACGAGACCTGAAAGAAAGAAATTTATAGACAGCTTTAGGTGAG 1179
 DB 1384 GTGCTAGGGGACGACTTTCTTCAACCGCAGCAGGAGCTGACGAGACCTTGCGATGAG 1443
 QY 1180 AAGATATTCAAGTACAGCCACTACAAATGAAAGATGAGATTTCCCAATGATATTGCA 1239
 DB 1444 AAGTATCATCCGATACACCTGTATCGGTGTTCACCCACGAGACAC---GACCTGTGC 1500
 QY 1240 TTGCTCAAGTTAAAGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 1299
 DB 1501 CTGATCGGGCTGAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560
 QY 1300 GTGCTGTGCTGCTGAT-----GGTCTCTTTCCTCTGTGAGATGATGATGATGATG 1353
 DB 1561 ATCTGCTGCGCCGAGCGCGGACGACCTTCCCGCAGACACAAAGTGCACAAATTTGCGGG 1620
 QY 1354 TGGGG-----TGTACAGAAACAGGAAAGAGGTCCCGCAGCTCTGATGATGATGATG 1407
 DB 1621 TGGGGCCACTGTGATGAGAACTGAGCGGCTATCTCAGCTCCCTGCGGAGGAGGCTGTC 1680
 QY 1408 AAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1467
 DB 1681 CCCGTGTGCGCCGACCAAGTGCAGACGCTGAGTCTAGGCGCGGACATCAAGCCCC 1740
 QY 1468 AGTATGATCTGTGAGAAATCTTCAAGAACTTGGCAAGACACCTGCGAGGATGATCT 1527
 DB 1741 AACATGCTGTGTGCGCGCTACTT---CGACTGCAAGTCCGACGCTTGGCAGGGGACTCA 1797
 QY 1528 GAGGCGCCCTGACCTGTGAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1587
 DB 1798 GGGGGGCGCCCTGTGCTGAGAAAGCGCGCTGATCTTACTCTAGGAGATCATGATGCG 1857
 QY 1588 GGCCTGAGATGTGG-----AAGAGCCAGAGGCTTACACCCAGTTACCAATTTCTG 1641
 DB 1858 GGTGAGGCTGCGGGGCGCTCCACAAGCGGGGCTTACACCGCGGTGCGCAACTATGTG 1917
 QY 1642 AATTGATCAAAAGCACCAT 1661
 DB 1918 GACTGTGATCAACGACCGAT 1937
 RESULT 15
 ABA70919/C
 ID ABA70919 standard; DNA: 117 BP.
 XX
 AC ABA70919;
 XX
 DT 01-FEB-2002 (first entry)
 XX
 DE Human foetal liver single exon nucleic acid probe #19224.

XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
 KW Homo sapiens.
 OS
 XX WO200157277-A2.
 PN
 XX 09-AUG-2001.
 PD
 XX 30-JAN-2001; 2001WO-US00669.
 PF
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483447/52.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 XX analyzing gene expression in human fetal liver -
 PS Claim 4; SEQ ID NO 19224; 639pp + sequence listing; English.
 XX
 CC The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC foetal liver. The present sequence is a single exon nucleic acid
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SO Sequence 117 BP; 24 A; 24 C; 40 G; 29 T; 0 other;

Query Match 7.0%; Score 117; DB 22; Length 117;
 Best Local Similarity 100.0%; Pred. No. 8e-24;
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 332 TGCATAATGTCGCAAGGACACCATGTCGCGGGCCATGTCATTTACCCAGAGTC 391
 |
 DB 117 TGCATAATGTCGCAAGGACACCATGTCGCGGGCCATGTCATTTACCCAGAGTC 58
 |

OY 392 CTCCTTACTACCGCTGTGTCTGTAAACACCCCTTACACAGTCCAGCTGCTCCAG 448
 |
 DB 57 CTCCTTACTACCGCTGTGTGTGTAAACACCCCTTACACAGTCCAGCTGCTCCAG 1
 |

Search completed: March 6, 2003, 16:01:34
 Job time : 353.5 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 15:20:13 ; Search time 2178.5 Seconds
(without alignments)
12511.822 Million cell updates/sec

Title: us-09-912-559-1

Perfect score: 1683
Sequence: 1 atgttgcagatgtctga.....aaagtgaagtgtctctaa 1683

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estinu:*
5: em_estrov:*
6: em_estropl:*
7: em_estro:*
8: em_hc:*
9: gb_estl:*
10: gb_estl2:*
11: gb_hc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	698.2	41.5	891	13	BI761782 603046775
2	451.6	28.8	592	10	BE032018 130809 MA
3	447.2	26.6	802	13	BC972579 602841165
4	437.8	26.0	674	9	AA217892 mv54b08.r
5	435.4	25.9	751	13	BI148082 602912432
6	435.2	25.9	570	13	BM508620 1137h03.y

Result No.	Score	Query Match	Length	DB ID	Description
7	416.4	24.7	929	12	BF788188 602113411
8	396.4	23.6	791	13	BI332440 602880841
9	395.8	23.5	546	9	AA237499 m10b10.r
10	392.6	23.3	608	10	AV601564 602105493
11	390.4	22.9	902	12	BF780971 602105493
12	386	22.9	982	13	BI220028 602934913
13	384.6	22.9	969	12	BF785781 602112402
14	372.4	22.1	654	12	BF384535 602046804
15	370.6	22.0	654	10	AMA75402 602046804
16	370.4	22.0	587	13	BC972681 602839029
17	365.6	21.7	659	10	AM610902 602839029
18	363.6	21.6	594	13	BM503097 602105493
19	348.8	20.7	716	12	BF789705 602103735
20	335.8	20.0	655	10	BB569555 602103735
21	330.8	19.7	934	13	BI765113 603051314
22	327.8	19.5	345	14	T68666 y43e09.r1
23	319	19.0	486	9	AA268125 602103735
24	311.8	18.5	600	12	BF788541 602114406
25	304	18.1	335	14	R89458 yq01a10.r1
26	293.2	17.4	304	14	T77362 yd72902.r1
27	279.8	16.6	521	13	BI221626 602936942
28	275.8	16.4	912	13	BC972626 602836922
29	275.6	16.4	366	14	T93666 y606a11.r1
30	266	15.8	434	14	T84369 yd37g04.r1
31	258.6	15.4	781	12	BF533788 602075308
32	254.6	15.1	458	14	R10295 yf36e12.r1
33	245.4	14.6	719	13	BI102281 602885594
34	242.8	14.4	593	9	A1035669 ud60b09.x
35	242.6	14.4	384	14	M66981 me29h02.r1
36	242.6	14.4	397	9	AA002504 mg45a03.r1
37	241.8	14.4	513	13	BI247736 602959379
38	241.4	14.3	353	14	R11014 yf38g10.r1
39	235.6	14.0	489	9	A1787533 u119e04.y
40	235.4	14.0	393	14	T50880 yb88g07.r1
41	215	12.8	530	10	BE235171 142539 MA
42	204.8	12.2	361	9	AA108230 mp38a11.r1
43	199.4	11.8	320	14	T69534 yd12f11.r1
44	154.6	9.2	642	13	BJ067550 BJ067550
45	150.4	8.9	810	9	A1255992 u176c01.x

ALIGNMENTS

RESULT 1
BI761782 891 bp mRNA linear EST 25-SEP-2001
LOCUS 603046775F1 NIH_MGC_116 Homo sapiens CDNA clone IMAGE:5187066 5',
DEFINITION mRNA sequence.
ACCESSION BI761782
VERSION BI761782.1 GI:15753360
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 891)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L14M11467 row e column: 19
High quality sequence stop: 839.
Location/Qualifiers 1..891

FEATURES
source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5187066"
/clone_id="NIH_MGC_116"
/lab_host="DH10B"
/Note="Organ: pooled colon, kidney, stomach; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH MGC library."

```

BASE COUNT 243 a 234 c 211 g 203 t

Query Match 41.5%; Score 698.2; DB 13; Length 891;
 Best Local Similarity 97.2%; Pred. No. 1.4e-187;
 Matches 775; Conservative 0; Mismatches 13; Indels 9; Gaps 6;

```

OY 1 ATGTTGGCCAGATGTCGATCTCATGTTCTGCTTAATGGCTGTGGGAAAGACA 60
DB 81 ATGTTGGCCAGATGTCGATCTCATGTTCTGCTTAATGGCTGTGGGAAAGACA 140
OY 61 GCCTGGGTTCTCCCTGATGTCCTTTATTGGAAGCTGGACCCAGACTGGACCCCTGAC 120
DB 141 GCCTGGGTTCTCCCTGATGTCCTTTATTGGAAGCTGGACCCAGACTGGACCCCTGAC 200
OY 121 CAGTATGATTACAGTACGAGAGATTATATCAGAGAACACACAGTACAGACCTTACC 180
DB 201 CAGTATGATTACAGTACGAGAGATTATATCAGAGAACACACAGTACAGACCTTACC 260
OY 181 CAGTATGATTACAGTACGAGAGATTATATCAGAGAACACACAGTACAGACCTTACC 240
DB 261 CAGTATGATTACAGTACGAGAGATTATATCAGAGAACACACAGTACAGACCTTACC 320
OY 241 CCCTGGGAACAGGTGGGAGTGCCTGCTCATGGAGAGACCTTACATGGAGTGGCTG 300
DB 321 CCCTGGGAACAGGTGGGAGTGCCTGCTCATGGAGAGACCTTACATGGAGTGGCTG 380
OY 301 GCTCTCTTCTCGGGAATAGTGTACAGAAATGCAAAATACGTGCAAGACCAACCATGT 360
DB 381 GCTCTCTTCTCGGGAATAGTGTACAGAAATGCAAAATACGTGCAAGACCAACCATGT 440
OY 361 GCGCGGGGCAATGTCATTAACCAAGAGTCGCTTACTACCGCTGTGTGTGAAGAC 420
DB 441 GCGCGGGGCAATGTCATTAACCAAGAGTCGCTTACTACCGCTGTGTGTGAAGAC 500
OY 421 CTTTACACAGTCCAGCTGCTCCCAAGTGGTCTGTATGAGGCAAAACCCCTGACAG 480
DB 501 CTTTACACAGTCCAGCTGCTCCCAAGTGGTCTGTATGAGGCAAAACCCCTGACAG 560
OY 481 AATGGGCTACTGCTCCCGCATTAACGGAGATCCAAATTAACCTGGCGCTGCCGAC 540
DB 561 AATGGGCTACTGCTCCCGCATTAACGGAGATCCAAATTAACCTGGCGCTGCCGAC 620
OY 541 CAGTTCAAGGGGAAATTTGTGAATATAGTTCATGATGACTGCTATGTTGGAGTGGTAC 600
DB 621 CAGTTCAAGGGGAAATTTGTGAATATAGTTCATGATGACTGCTATGTTGGAGTGGTAC 680
OY 601 TCTTACCGAGGAAATATAGTACAGTTC -AACCGAGATGGTGGCTTTATCGAAGTTC 659
DB 681 TCTTACCGAGGAAATATAGTACAGTTC -AACCGAGATGGTGGCTTTATCGAAGTTC 740
OY 660 CC-ACCTCTCTTGGAGAGATTAACAATGTTTATGAGGAG - -ATGCTGAAGAACCATG 715
DB 741 CCAAGCTCTCTTGGAGAGATTAACAATGTTTATGAGGAGATGCTTGAAGAACCATG 800
OY 716 GGAATTGGGAGAA-CACATTTCTGCAAG-AAACCCAGATCGGAGG-AAAACCCCTGGTGC 771

```

```

DB 801 GGAATTGGGAGAAACACCAATTTCTGCAAGAAACCCAGATGCCGAGAAAAAGCCCTGTGC 860
OY 772 TTTATTAAGTTTACCAA 788
DB 861 TTTATTAAGTTTACCAA 877

```

```

RESULT 2
BE032018 592 bp mRNA linear EST 09-JUL-2000
LOCUS BE032018
DEFINITION 130809 MARC 1Pig Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BE032018
VERSION BE032018.1 GI:8327027
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

```

```

REFERENCE
1 (bases 1 to 592)
Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.,
and Keeler,J.W.
Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
Unpublished (2000)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt-trimmed with phred
v0.980904.e. Vector identified by cross-match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCACAGTACAGAG
Plate: 69 row: K column: 12
Seq primer: ATTGAGTACATATG.
location/qualifiers
1..592
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_id="MARC 1Pig"
/tissue_type="pooled"
/lab_host="DH10B"
/Note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

```

FEATURES
 source

BASE COUNT 144 a 152 c 174 g 122 t

Query Match 26.8%; Score 451.6; DB 10; Length 592;
 Best Local Similarity 86.0%; Pred. No. 2.2e-117;
 Matches 518; Conservative 0; Mismatches 69; Indels 15; Gaps 1;

```

OY 522 CACCTGTCGTCGTCGCCAGCACTTCAAGGGAATTTCTGGAATATAGTTCGATGACTG 581
DB 6 CCGGTGGGCTGTCTGACCACTTTAAGGGAGTTCTGTGAATATAGTTCGATGACTG 65
OY 582 CTATGTTGGGAGTGGCTACTCTTACCGAGGGAATATAGTACAGTCAACAGCATGC 641
DB 66 CTATGTTGGGAGTGGCTACTCTTACCGAGGGAATATAGTACAGTCAACAGCATGC 125
OY 642 GTGCTTACTTGAACATCCCACTCTCTTGCAGAGAAATTAACAGTGTATGAGAGA 701
DB 126 GTGCTTACTTGAACATCCCACTCTCTTGCAGAGAAATTAACAGTGTATGAGAGA 185
OY 702 TGCAGAAACCATGAGGATTTGGGAGAACATTTCTGAGAAACCCAGATGCGAGAAAA 761
DB 186 TGCAGAGCCCATGAGGATTTGGGAGAACATTTCTGAGAAACCCAGATGCGAGAAAA 245
OY 762 GCCCTGTGCTTTATTAAAGTTTACCAATGACAGGTGAATGGGAATGATGATGCTC 821

```

D	b		 GGCGTGTGTTTCAATTAAAGTAGACAGTGCCAACTGAATGGAGAGTACTGTGACGTCCC	305
Oy	822	AGCCTGTCTCACCCAGSAGCTTGCTCTACCACAGAGAAAGCCCTACTGAGCATTAACCAA	861	
D	b	TGCCTGTCTCACAC-----CCAGAGAGGAAAGCCCTTAAGAAGCTGTGACCGA	350	
Oy	882	GCCTCCGGGGTTTTGACTCTGTGTGAAGAACTGAGATAGTCAGAGAGAGATCAAGACAAT	941	
D	b	 GGTTCGCCGGTTTACTCTGTGTGGAGGACAGACAGCAGAAAGAGTTCAAGAGAGAT	410	
Oy	942	CCTATGAGAGCTTTAAAGACACGGCGGGGACAGACCCTATGGCAGAGCGTCCCTCCAGTCTC	1001	
D	b	 CTACAGAGAGCTTTAAGACACGGCGGGGACACACCCTGTGGCAGCGTCCCTACAGAGCTC	470	
Oy	1002	GCTGCTCTGACACTCTCCATGCCCCAGGAGGACCTTGTGTGTGGGGGCGCTGATCCACCC	1061	
D	b	 CTTGCACTGACCGCTCTCCATGCCCCAGGAGGACCTTGTGTGGGGGCGCGTATCCACCC	520	
Oy	1062	CTGCTGGGTGCTCACCTCTGTCCCACTGACACCGACATAAAGACAGACATCTTAAGGTGT	1121	
D	b	 CTGTGTGGGTGTGACCGCTCCCATTTGACACGACATAAAGCCCAATATCTTAAGATGT	590	
Oy	1122	GC 1123		
D	b	591 GC 592		
R	E	S		
R	E	S		
L	O	C		
D	E	F		
A	C	C		
V	E	R		
S	I	O		
N	E	F		
I	N	I		
O	N	I		
R	G	A		
N	A	M		
C	O	M		
M	E	N		
E	N	T		
R	E	F		
E	N	T		
J	O	U		
R	N	A		
N	A	L		
C	O	M		
M	E	N		
E	N	T		
J	O	U		
R	N	A		
N	A	L		
C	O	M		
M	E	N		
E	N	T		
J	O	U		
R	N	A		
N	A	L		
C	O	M		
M	E	N		
E	N	T		
J	O	U		
R	N	A		
N	A	L		
C	O	M		
M	E	N		
E	N	T		
J	O	U		
R	N	A		
N	A	L		
C	O	M		
M	E	N		
E	N	T		
J	O	U		
R	N	A		
N	A	L		
C	O	M		
M	E	N		
E	N	T		
J	O	U		
R	N	A		
N	A	L		
C	O	M		
M	E	N		
E	N	T		
J	O	U		
R	N	A		
N	A	L		
C	O	M		
M	E	N		
E	N	T		
J	O	U		
R	N	A		
N	A	L		
C	O	M		
M	E	N		
E	N	T		
J	O	U		
R	N	A		
N	A	L		
C	O	M		
M	E	N		
E	N	T		
J	O	U		
R	N	A		
N	A	L		
C	O	M		
M	E	N		
E	N	T		
J	O	U		
R	N	A		
N	A	L		
C	O	M		
M	E	N		
E	N	T		
J	O	U		
R	N	A		
N	A	L		
C	O	M		
M	E	N		
E	N	T		
J	O	U		
R	N	A		
N	A	L		
C	O	M		
M	E	N		
E	N	T		
J	O	U		
R	N	A		
N	A	L		
C	O	M		
M	E	N		
E	N	T		
J	O	U		
R	N	A		
N	A	L		
C	O	M		
M	E	N		
E	N	T		
J	O	U		
R	N	A		
N	A	L		
C	O	M		

[illegible]


```

Db 363 AGTATGATTTTGGGGGAGACCTTCAGAAAGCCCGGATCAGACACCTCCAGGCTGATCG 422
QY 1528 GGAGGCCCTGACCTGTGAGAGAGAGCGCCTACTACTGTATGGATAGTGAAGCTGG 1587
Db 423 GGGGGCCCTCAACCTGTGAGAGAGATGGAACTTACTACCTGCGGAGATTGAGCTGG 482
QY 1588 GGCTGTGAGTGGGAGAGAGCGGAGTCTACACCCAGTACCAATTCCTGATTTGG 1647
Db 483 GGGCAGAGATGTGGAGAGAGCCGAGGATCTACACTCAAGTACCAAGTTCCTGATTTGG 542
QY 1648 ATCAAGCCACCATCAAAAGTGAAGTGTG 1675
Db 543 ATAAAGACCCATGACAGGGAGGCTGG 570

RESULT 7
LOCUS Bf788188 929 bp mRNA linear EST 12-JAN-2001
DEFINITION 60211341JF1 NCI_CGAP_Ki414 Mus musculus cDNA clone IMAGE:4241642
5', mRNA sequence.
ACCESSION Bf788188
VERSION Bf788188.1 GI:12093224
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
1 (bases 1 to 929)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://lmage.llnl.gov
Plate: L14M9858 row: e column: 03
High quality sequence stop: 686.
Location/Qualifiers
1..929
source
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_1db="IMAGE:4241642"
/lab host="NCI_CGAP_Ki414"
/note="Organ: Kidney; Vector: PCMV-SPORT6; Site: 1; NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Constructed by Life
Technologies. Note: this is a NCI/CGAP Library. 1"

BASE COUNT 238 a 275 c 253 g 162 t 1 others
ORIGIN
Query Match 24.7%; Score 416.4; DB 12; Length 929;
Best Local Similarity 75.9%; Pred. No. 3e-107;
Matches 529; Conservative 0; Mismatches 162; Indels 6; Gaps 1;
QY 1 ATGTTGCCAGAGATGCTGATCTCATGTCCTGTTAATGCTGTGGTGAAGACAA 60
Db 44 ATATTTGTCAGATGATGTTCCGTCCTGCTGCTGCTAATGCTGGGGAAGTGA 103
QY 61 GCCTGGGCTTCCCTGATGCTTTTATGAAAGCTTGACCCAGACTGGACCCCTGAC 120
Db 104 GTATGGGCTCTCATGATGTCCTTCATGCGCCGCCAGACAGATNGACCCCGCAT 163
QY 121 CAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
Db 164 GACTATTACTACAGCTATGAGACATCCAGCCGACAGACAGACCCAGTGTACGCAAGCC 223

```

```

QY 181 CATGCTGAGAAATCTCTGATCTACTACATGAGAGCCAGCTGATCATCCAGGCCAAC 240
Db 224 ACCCTCGAGAAACCCGAGCTGTACT-----ATGAGACATATATCATGACGATCCAC 277
QY 241 CCCTGTGAACAGGCTGGGAGCTGCTGTCATGAGAGACCTTTCATGATGCTGCTG 300
Db 278 CCCTGTGAACAGGCTGGGAGCTGCTGTCATGAGAGAGGATGATCTGAGTTGCGAGGCCA 337
QY 301 GCCTCTTCTCTGGGAAATGATGTCAGAAAGTGCAGAAATACGTCAAGAGCAACCATGT 360
Db 338 GCGCCCTTCTCTGGGAGAGCGGCTGCCAGACTGACAAAACAGTGAAGAGCAACCATGT 397
QY 361 GCGCGGGGCAATGTCATTTACCCAGAGACTCTCTCTACTACCGGCTGTCTGTAACAC 420
Db 398 GTCCATGATGATGTCCTCATTTACCCAGAGACCCCTTACTACCGCTGTCTGTAACAC 457
QY 421 CCTTACAGAGGTCCAGCTGCTCCCAAGTGTCTCTGTATGACAGGCAACCCCTGCCAG 480
Db 458 CCTTACAGGAGACAGACTGCTCCAAAGTGTCTCCGCGATGACAGGCCAAACCCCTGCCAG 517
QY 481 AATGGGCGTACCTGCTCCGCGATGAGAGATGCAAGTTCACTGTGCTGTCCGAC 540
Db 518 AATGGGCGAGTCTGTCTCCACACAGAGAGATGCAAGTTTACTGTGCTGTCCGAC 577
QY 541 CAGTTCAAGGGAATTTCTGTGAATAGTTCGATGACCTGTATGTTGGCATGCTAC 600
Db 578 CAGTATTAAGGGGAATTTCTGTGAATAGTTCGAGAGCTGTATGTTGGTATGCTAC 637
QY 601 TCTTACGAGGGAATTAATGATGACATGACATGACATGACATGACATGACATGAC 660
Db 638 TCTTACGAGGGAATTAATGATGACATGACATGACATGACATGACATGACATGAC 697
QY 661 CAGCTCTCTTTCAGAGAGATTTACAACTGTTTATGG 697
Db 698 CAGCTCTCTTTCAGAGAGATTTACAACTGTTTATGG 734

RESULT 8
LOCUS B1332440 791 bp mRNA linear EST 30-JUL-2001
DEFINITION 60298084JF1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:5133587 5',
mRNA sequence.
ACCESSION B1332440
VERSION B1332440.1 GI:15017097
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
1 (bases 1 to 791)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://lmage.llnl.gov
Plate: L14M1328 row: a column: 12
High quality sequence stop: 787.
Location/Qualifiers
1..791
source
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_1db="IMAGE:5133587"
/lab host="NCI_CGAP_L19"
/note="Organ: Liver; Vector: PCMV-SPORT6; Site: 1; NotI;

```


Site 2: SalI: Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.9 kb. Constructed by Life
 Technologies. Note: this is a NCI-CGAP library."

BASE COUNT 202 a 234 c 194 g 161 t
 ORIGIN

Query Match 23.6%; Score 396.4; DB 13; Length 791;
 Best Local Similarity 75.6%; Pred. No. 1.4e-101;
 Matches 520; Conservative 0; Mismatches 161; Indels 7; Gaps 2;

OY 1 ATGTTGCGAGATGCTGATCCATGTCGCTGTTATGCTGGGGAAGCA 60
 || ||||| ||||| || ||||| ||||| ||||| ||||| ||||| |||||
 Db 23 AATATTGACAGATGTGGTGTCCGTGCTGCTATGCCCTGGGGAAGCA 82
 OY 61 GCCTGGGTTCTCCCTGATGTTTATTGAAAGCTGACCCAGACTGAC 120
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 83 GTCATTGGGCTCCTCAGATGCTCTCATGCGCCAGAGCCAGATTGAC 142
 OY 121 CAGTATGATTACAGTACGAGATTTATATGAGAGAGACACAGTAC 180
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 143 GACTATTACTACATGATGAGCTGACAGCCAGAGACCCAGTGTAC 202
 OY 181 CATGTGAGATCTCTGACTACTACACTGAGAGACCACTGATCCAG 240
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 203 ACCCTGAGAACCCGACGTGTACT-----ATGAGAGCATGATCTC 256
 OY 241 CCCTGTGACACGCTGGGAGCTGCTCCATGAGACACCTTACATGAC 300
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 257 CCTGTGACACGCGGGGAGCTGATCATCAGAGGAGATACCTTCA 316
 OY 301 GCTCTTCTCTGGGATTAAGTGCAGAAATGCAAAATACGTGCAAGAC 360
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 317 GCCCCTCTCGGGGAGCCGCTGACACTGCACAAACAAAGTCAAGAC 376
 OY 361 GCGCGGGGCAATGCTCATTAACCAAGTCTCTTACCGCTGTCTGTA 420
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 377 GTCATGCTGATGCTCTCATTAACCAAGACCTCTTACCGCTGTCT 436
 OY 421 CCTTACACAGTCTCCAGCTCTCCCAAGTGTCTGTATGACAGCCCA 480
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 437 CCTTACACAGGAGACGATCTCTCAAAAGTGTCTGCGCATGAGCA 496
 OY 481 AATGAGGCTTACCTCTCCGCAATAGCAGAGATCAAGTTCACTGCT 540
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 497 ATGGCGGAGTCTGTCCGACACAGCAGAGATCCAGTTTACTGTCT 556
 OY 541 CAGTTCAGGAGAAATCTGTGAA-ATAGTCTGATGATGATGATGGA 599
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 557 CAGTATTAAGGGAATTTCTGTGACATAGTCTGACACACTGATG 616
 OY 600 CTCTTACGAGGAAATGATAGACACTCAACCAAGATGCTGCTT 659
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 617 CTCTTACGAGGAAATGATAGACACTCAACCAAGATGCTGCTG 676
 OY 660 CCAGCTCTCTTGCAGGAGATTAAC 687
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 677 CCAGCTCTCTTGCAGGAGATTAAC 704

RESULT 9

AA237499 546 bp mRNA linear EST 03-MAR-1997
 LOCUS mx10b10.1 Soares mouse NML Mus musculus cDNA IMAGE:679771 5'
 DEFINITION similar to TR:G1345398 G1345398 HGF ACTIVATOR LIKE PROTEIN. 1, mRNA
 sequence.

ACCESSION AA237499
 VERSION AA237499.1 GI:1861520
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 546)
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS

Matta, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Giesel, S., Kneba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Stepien, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.

THE WASHU-HM1 MOUSE EST PROJECT
 UNPUBLISHED (1996)
 CONTACT: MATT, M/MOUSE EST PROJECT
 WASHU-HM1 MOUSE EST PROJECT
 WASHINGTON UNIVERSITY SCHOOL OF MEDICINE
 4444 FOREST PARK PARKWAY, BOX 8501, ST. LOUIS, MO 63108
 TEL: 314 286 1800
 FAX: 314 286 1810

EMAIL: mouseest@wustl.edu
 THIS CLONE IS AVAILABLE ROYALTY-FREE THROUGH LBNL. CONTACT THE
 IMAGE CONSORTIUM (INFO@IMAGE.LBNL.GOV) FOR FURTHER INFORMATION.
 MGI:419475

POSSIBLE REVERSED CLONE: SIMILARITY ON WRONG STRAND
 Seq primer: -28m13 rev2 ET from Amersham
 High quality sequence stop: 405.
 Location/Qualifiers

FEATURES

Source

1. 546
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:679771"
 /clone_lib="Soares mouse NML"
 /lab_host="Liver"
 /lab_type="Liver"
 /note="vector: pT73D-Pac (pharmacia) with a modified
 polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer (5',
 TGTACCAATCTGTAAGTGGAGCGGCCGCAATCTTCTTTTCTTTTCTT
 3');
 double-stranded cDNA was ligated to Eco RI adaptors
 (pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 constructed and normalized by Bento Soares and M. Fatima
 Bonaldo."

BASE COUNT 141 a 142 c 151 g 112 t
 ORIGIN

Query Match 23.5%; Score 395.8; DB 9; Length 546;
 Best Local Similarity 83.1%; Pred. No. 1.7e-101;
 Matches 451; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

OY 868 GAGCATCAACCAAGTCTCCGGGTTTGTACTCTGTGGAAGAGTGAATGACAG 927
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 4 GAGCTGTGATGAGATGCTGACAGGTTCAGATCTGAGAGAAAGAGTGAAC 63
 OY 928 AAGATCAAGAGATCTATGAGAGCTTTTAAGACAGCGGCGCAACAC 987
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 64 GCAATCAAGCTATCTAAGGAGGCTTTAAGACACAGACAGACACCGTGG 123
 OY 988 TCCCTCAGTCTGCTGCTCTGACATCTCCATGCCCCAGGGCCACTTCTG 1047
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 124 TCCCTCAGACCTTCACTGCTGCTGACATCTCCATGCCCCAGGGCCACT 183
 OY 1048 GCGGTATCCACCCCGTGGGGTGTCTACCTGCTCCAGTGCAGACATTA 1107
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 184 GCGGTATCCACCCCGTGGGGTGTCTACCTGCTCCAGTGCAGACATTA 243
 OY 1108 CATCTAAGATGCTGTAGGAGACAGACTGAGAGAGAAATTTTCATGAG 1167
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 244 CATCTAAGATGCTGTAGGAGATCAGAGCTGAGAGAGAAATTTTCATGAG 303
 OY 1168 TTTAGGTGAGAGATATTCAAGTACAGCACTACATGAGAGATGAGATT 1227
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 304 TTTAGGTGAGAGATATTCAAGTACAGCACTACATGAGAGATGAGATT 363
 OY 1228 AATGATGATGATGCTCAAGTTAAGCAGTGAATGCTGCTGATTCACAA 1287
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 364 AATGATGATGCTGCTCAAGTTAAGCAGTGAATGCTGCTGATTCACAA 423
 OY 1288 TAGGTGAGAGATGCTGCTGCTGATGAGTCTTCTTCTGAGAGTGGCAGATC 1347

Fri Mar 7 09:21:10 2003

us-09-912-559-1.rst

Page 8

Accession	Sequence	Position
Db	TATGTAAACATGTAATTTTGCCACGACCCCTTTCCCTCTGAACATAAGGCCACATC	483
Db	424	
QY	TCTGCTGGGTGGTGTACAGAAACAGAAAAAGGGTCCGCCAGCTCCTGTGATGCCAATC	1407
QY	1348	
Db	TCGTGGTGGGGGTGTACAGAAACAGGAGGGTTCCGCCAGCTCCTGTGATGCTAAATC	543
Db	484	
QY	1408 AAC 1410	
QY	111	
Db	544 AAG 546	

RESULT 10	AV601564	608 bp	mrna	linear	EST 27-NOV-2001
LOCUS	AV601564				
DEFINITION	AV601564 Bos taurus kidney fetus cdna clone E1K1005A04				
ACCESSION	5', mRNA sequence.				
VERSION	AV601564				
KEYWORDS	AV601564.1 GI:9723884				
SOURCE	EST.				
ORGANISM	COW, <i>Bos taurus</i>				

REFERENCE
TAKASUJI, A., HIROTSUNE, S., ITOH, R., JIHOZONO, A., SUZUKI, H., ASO, H.
1 (bases 1 to 608)

TITLE	JOURNAL
Establishment of a high throughput EST sequencing system using poly(A) tail-removed cDNA libraries and determination of 36,000 bovine ESTs	Nucleic Acids Res. 29 (22), E108 (2001)
2102054	

COMMENT
Contact: Yoshikazu Sugimoto
Animal Genetics Division
Shirakawa Institute of Animal Genetics
Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
Tel.: 81-248-25-5641
Fax: 81-248-25-5725
Email: kazusugi@ccocn.ocn.ne.jp
Single pass sequencing. This clone was obtained from a polyA-deleted cDNA library.

FEATURES	source
Location/Qualifiers	1. .608
/organism="Bos taurus"	
/db_xref="taxon:9913"	
/clone="EIKT005A04"	
/clone_1b="Bos taurus kidney fetus"	
/tissue_type="kidney"	
/dev_stage="fetus"	
/lab_host="DH10B"	
/note="Vector: pZ11, Site_1: SalI; Site_2: NotI; Poly A was deleted from a NotI site"	
148 a 171 c 151 g 135 t 3 others	

Query Match	23.3%	Score 392.6;	DB 10;	length 608;
Best Local Similarity	81.9%;	Pred. No. 1.5e-100;		
Matches 465; Conservative	0;	Mismatches 97;	Indels 6;	Gaps 1

QY	1	ATGTTTGCCAGAGATGTCTGATCTCCATGTTCTTCTGCTGTTAAATGGCTCTGGTGGGAAAGACA	60
Db	47	ATGTTTGCCAGAGATGTCGATCTCCATGTTCTTCTGCTGTTAAATGGCTCTGGCGGGAAAGACA	106
QY	61	GCGCTGGGGTTCCTCCGATGTCTTTATTGGAAACCTGGAGCCAGACGTGAGACCCCTTAC	120
Db	107	GCGTTGGGGCTCTCCCTGCTCTGCTCTTTCTTCACAGAGCCGAGACCTGAGTTGGACTCTTAC	166
QY	121	CAGTATGATTACAGCTACGAGAGATTATAATCAGSAAAGAAACAACGTAAGCACTTACC	180
Db	167	CAGTATGAGTACAGCCAGAGATATAATACCCAGSAAAGAAACGCGAGTACCACTGCC	226
QY	181	CATGCTGAGAAATCCCTGACTGTGTACTACACTGAGGAGCAAGCTGATCCATGCGCCAGCCCAAC	240

Db	227	TATTTGCAACACCTGTACGTGACTACGAAAGAGA-----TGACCCACGCTGTCCAC	280
Qy	241	CCCTGTGAACACGGTGGGGACGCTCTCCATCGAGACACCTTCACATGACGTGGCGT	300
Db	281	CCCTTCACATGTGTGGGGAGTCCCTCTCAGTGGGGACACCTTCACGTCCGCTTGC	340
Qy	301	GCCTCTTCTGTGGAAATAGTGTGAGAAATGTGCAAAATACGTGCAAGACCAACCCATGT	360
Db	341	GACCTTTCTGTGAAACAGGTGTGAGATGTGCAAAACAGTGCAAAACACCCCTGT	400
Qy	361	GGCGGGGGCCATGTCTCATTTACCCAGAGTCTTCCCTACTACCGCTGTGTGTAAAC	420
Db	401	GGCGGGGAGACTGTCTCATTTACTGTAGAGTCTCTTACCCACGGCTGTGCTGAAACAC	460
Qy	421	CGTTACAGAGTCCCAAGTGTCTGCCAAGTGTTCCTGTATGACAGCCAAACCCCTGCCAG	480
Db	461	CCTTACAGGGGTTTACAGCTGTTCAGAGTGTGCTGTCTGTGACAGGCCAAATCCCTGACAA	520
Qy	481	AATGGGGTACTCTCTCCGGATTAAGGGAGGATTCCAAGTCACTGTGCTGTGCCGAC	540
Db	521	AAGGGTACCTCTCTCCGGCAGAGAGGGAGGTCCAAAGTTCACCTGTGCTGTGCAC	580
Qy	541	CAGTTCAAGGGAAATTTCTGAATAG	568
Db	581	CAGTTCAAGGGAAAGCTTTTGGAAATAG	608

RESULT 11	
LOCUS	BF780971
DEFINITION	BF780971 902 bp mRNA linear EST 12-Jan-2001
ACCESSION	602105493F1 NCI-CGAP_kid14 Mus musculus cDNA clone IMAGE:4223710
VERSION	5', mRNA sequence.
KEYWORDS	BF780971 BF780971.1 GI:12086004
SOURCE	EST.
ORGANISM	house mouse. Mus musculus

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.	NIH-MGC	http://mgc.nci.nih.gov/		
(bases 1 to 907)				
		National Institutes of Health, Mammalian Gene Collection (MGC)		
		Unpublished. (1999)		
Contact:	Robert Strausberg,	Ph.D.		

Email: Ggaps-f@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>
Plate: LIA9812 row: 1 column: 23
High quality sequence stop: 117.

FEATURES	source
location/Qualifiers	1..902
	/organism="Mus musculus"
	/strain="FVB/N"
	/db_xref="taxon:10090"
	/clone_image="4223710"
	/clone_lib="NCI-CGAP Kid14"
	/lab_host="DH10B (TI phage-resistant)"
	/note="Organ: kidney; Vector: pCMV-Sport6; Site.1: NotI; Site.2: SalI; Cloned unidirectionally. Primer: oligo dr. Average insert size 1.75 kb. Constructed by life Technologies. Note: this is a NCI-CGAP Library.]"
BASE COUNT	223 a 264 c 254 g 161 t
ORIGIN	

Query Match	23.2%;	Score 390.4;	DB 12;	Length 902;
Best Local Similarity	73.18;	Pred. No. 7.7e-100;		
Matches 532; Conservative	0;	Mismatches 186;	Indels 10;	Gaps 2;

```

QY 1 ATGTTTGCCAGATGTCGTATCTCCATGTCCTGTTAATGAGCTCTGAGGAAACACA 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 52 ATATTGTGACAGATGTTGGTGTCCGTCTCTGCTGATATGCCCCCTGTGGGAAATCA 111
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 GCGTGTGGGTTCTCCGATGTCCTTATTTATGGAAGCCCTGAGACCAAGCTCCGAC 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 112 GTCATGTGGCTCTCATGATGTCCTTATGCGCCCGACAGCCAGATTGACCCCGAT 171
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 CAGTATGTTACAGTACAGATATATATAGAGAAAGAACCAACAGTACACTTAC 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 172 GACTATCTACAGCTATAGACATCCAGCCAGACAGAACCCAGTGTGAGAGACC 231
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 CAGTGTGAATCTCTACTGCTACTACAGGACCAAGCTGATCCATGCGCAGCCAAC 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 232 ACCCGTGAACCCCGACGCTGACT-----ATCAACACATGATCCATGCCAGTCAAC 285
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 CCGTGTGAACAGGAGTGTGCTGCTCCATGAGACACCTTACATGAGAGTGTGCTG 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 286 CCGTGTGAACAGGAGGAGTGTATCATCAGAGGAGTACCTTCAAGTGTGAGCTGCCA 345
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 GCTCTTCTCTGAGATAAGTGTCAAGATGCAAAATGCAATGCGAGAGACACCATGT 360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 346 GCCCGCTTCTGAGGAGCCGCTGACACTCAGACAAACAGTGTGAGAGACACCATGT 405
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 GCGCGGGCCCATGTCTCATTAACCAAGTCTCTCTACTACCGCTGTGTGTAAACAC 420
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 406 GTCCATGTGTATGCTCTATACCCAGAACACCCCTACTACCGCTGTGTGTAAATAC 465
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 421 CCTTACAGAGTGTCCAGCTGTCCCAAGTGTCTGTATGAGGCAAAACCCCTGACAG 480
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 466 CCTTACAGAGGAGCAGACTGTCTCAAAAGTGTCTCGCATGAGGCAAAACCCCTGACAG 525
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 481 AATGGGGGTACTGCTGCTCCGAGATTAAGGAGATCAACCTGAGTGTGCTGCTGCTGAC 540
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 526 AATGGCGAGTGTCTGCTCCGACACAGACGAGATTCAGGTTTACCTGTGCTGCTGAC 585
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 541 CAGTTCAAGGGGAATTTCTGTAATAGCTCTGTATGATGCTGTATGCGATGCTTAC 600
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 586 CAGTATTAAGGGGAATTTCTGTAATAGCTCTGAGACAGCTGTATGCTGTGATGCTTAC 645
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 601 TCTTACAGGAGGAATTAATAGACAGTCAACAGCAGTCCGCTTACTTACTGAACTCC 660
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 646 TCTTACAGGAGGAATTAATAGACAGTCAACAGCAGTCCGCTTACTTACTGAACTCC 705
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 661 CACCTCTCTCTGAGGAGATTAACAACATGTTTATGAGAGATCGTGAACCCATGGGATT 720
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 706 CTCTCTCTCTGAGGAGATTAACAACATGTTTATGAGAGATCGTGAACCCATGGGATT 761
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 721 GGGGAACA 728
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 762 GGGGAACA 769
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

FEATURES
    Source
        1..682
        /organism="Mus musculus"
        /strain="FVB/N"
        /db_xref="taxon:10090"
        /clone="IMAGE:5098442"
        /lab_host="DH10B (T1 phage-resistant)"
        /note="Organ: liver; Vector: pCMV-Sport6; Site: 1; NotI;
        Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dT.
        Average insert size 1.9 kb. Constructed by Life
        Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT
    166 a 206 c 169 g 141 t
ORIGIN
    1 ATGTTTGCCAGATGTCGTATCTCCATGTCCTGTTAATGAGCTCTGAGGAAACACA 60
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    24 ATATTGTGACAGATGTTGGTGTCCGTCTCTGCTGATATGCCCCCTGTGGGAAATCA 83
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    61 GCGTGTGGGTTCTCCGATGTCCTTATTTATGGAAGCCCTGAGACCAAGCTCCGAC 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    84 GTCATGTGGCTCTCATGATGTCCTTATGCGCCCGACAGCCAGATTGACCCCGAT 143
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    121 CAGTATGTTACAGTACAGATATATATAGAGAAAGAACCAACAGTACACTTAC 180
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    144 GACTATCTACAGCTATAGACATCCAGCCAGAGAAACAGTGTGAGAGACACCATGT 203
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    181 CAGTGTGAATCTCTACTGCTACTACAGGACCAAGCTGATCCATGCGCAGCCAAC 240
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    204 ACCCGTGAACCCCGACGCTGACT-----ATCAACACATGATCCATGCCAGTCAAC 257
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    241 CCGTGTGAACAGGAGTGTGCTGCTCCATGAGGAGACCTTACATGAGTGTGCTG 300
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    258 CCGTGTGAACAGGAGGAGTGTATCATCAGAGGAGTACCTTCAAGTGTGAGCTCCA 317
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    301 GCTCTTCTCTGAGATTAAGTGTCAAGAAAGTCAAAATAGCTGTGACAGCAACCATGT 360
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    318 GCCCGCTTCTGAGGAGCCGCTGACAGTGCACAAACAAAGTGTGAGAGCAACCATGT 377
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    361 GCGCGGGCCCATGTCTCATTAACCAAGTCTCTCTACTACCGCTGTGTGTAAACAC 420
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    378 GTCCATGTGTATGCTCTCATTAACCAAGACACCCCTACTACCGCTGTGCTCAAAATAC 437
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    421 CCTTACAGAGTGTCCAGCTGCTCCATGAGTGTCTGTATGAGGCAAAACCCCTGACAG 480
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    438 CTTTACAGGAGACCAAGCTGTCCAAAGTCTTCCGCGATGAGAGCC-MACTCTGACAG 496
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    481 AATGGGCTACTGCTCTCCGCGATTAAGCGAGATCCAAGTACCTGTGCTGCTGAC 540
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    497 AATGGCGAGTGTCTGCTCCGACAGAGAGATCCAGGTATTAAGTGTGCTGCTGACAG 556
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    541 CAGTTCAAGGGGAATTTCTGTAATAGCTCTGATGACGTGATGTGCGATGCTTAC 600
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    557 CAGTATTAAGGGGAATTTCTGTAATAGCTCTGATGACGTGATGTGCGATGCTTAC 616
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    601 TCTTACAGGAGGAATTAAGAGAGTCAACAGCTGTGAGGAGCTTACTGAGAGTAC 659
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    617 TCTTACAGGAGGAATTAAGAGAGTCAACAGCTGTGAGGAGCTTACTGAGAGTAC 676
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    660 CCAC 663

```

RESULT 12
 LOCUS B1220028
 DEFINITION 602934913F1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:5098442 5',
 mRNA sequence.
 ACCESSION B1220028
 VERSION B1220028.1 GI:14673472
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.

Db	184	GACTATTACTACACTATATGAGCAGTCAGGCCACAGCAAGACCCCACTGTCCACGACAC	243
Qy	181	CAMGTGAGAAATCGTCAGTGTACTACATCAGACCAACAGTCATCTATGCAACCCAC	240
Db	244	ACCCCTGAGAAACCCCACTGTACT-----ATTAAGACGATGATTCATGCGCACTCCAC	297
Qy	241	CCCTGTGAACACGGTGGGAGACGCTCGCTCCATGGGAGACCTTCACATGAGTGGCTG	300
Db	298	CCGTGTGAACACGGCGGGAGCTGTATCATCAGAGGGGATACCTCATGTTGAGGCTGGCCA	357
Qy	301	GCTCTTCTCTGTGGAAATAAGTGTAGAAAGTGCAAAATACGTGCACAGACACACCATGT	360
Db	358	GCCCCCTTCTGGGGAGCCGGTGCACAGACTGCACAAAACAAGTGGACAAGCAACCCATGT	417
Qy	361	GGCGGGGGCCATCTTCATTTACCCAGAGTCTCTCCCTACTACCGCTGTGTGTGTAACAC	420
Db	418	GTCATGTGTATTCCCTCATTTACCCAGAGAACCCCTACTACCGCTGTGCTGCATAATAC	477
Qy	421	CCTTACACAGTCCACAGTGCATCCCAAGTGGTTCCTGTATCGAGGCGCAAAACCCCTGCAG	480
Db	478	CCTTACACGGGACACAGCTGTCCAAAGTGCTTCGGCATGAGGCGCAAAACCCCTGCCAG	537
Qy	481	AATGGGGCTACTCTCTCCGGCATTAAGGGAGATCCAAAGTTCACCTCGTGGCCCTGCCAG	540
Db	538	AATGGCGGAGTCTGTTCGACACACAGAGGGAGATCCAGGTTTAACTGTGCTGTGCCAGAC	597
Qy	541	CAGTTCAAGGGGAAATCTGTGAATAGTTCGTATGACTGCTATGTGTGGCGATGGCTAC	600
Db	598	CAGTTCAAGGGGAAATCTGTGAATAGTTCGAGACACACTGTTATGTCCGTATGGCTAC	657
Qy	601	TCTTACGAGGGAATGAATAGACAGTCAAC	633
Db	658	TCTTACCGA-GGCAAGTGAAGTAACACTCAAC	689

RESULT	15
LOCUS	AM475402
DEFINITION	654 bp mRNA linear EST 24-FEB-2000
ACCESSION	U985H07.1 Sugenone mouse kidney mKia Mus musculus cDNA clone
VERSION	IMAGEJ:2035349.5 similar to TR:Q14520 Q14520 HGF ACTIVATOR LIKE PROTEIN ; , mRNA sequence.
KEYWORDS	AM475402 AM475402.1 GI:7045508
SOURCE	EST.
ORGANISM	Mus musculus. house mouse.
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. I (bases 1 to 654)
AUTHORS	Maria,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person .B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurr,R., Ritter .E., Kohm,S., Shih,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson.R.
TITLE	The WashU-NCI Mouse EST Project 1999
JOURNAL	Unpublished (1999)
COMMENT	Contact: Maria W/WashU-NCI Mouse EST Project 1999 Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel.: 314 286 1800 Fax: 314 286 1810 Email: mouseest@wustl.edu This clone is available royalty-free through LNL; contact the IMAE Consortium (info@image.lnl.gov) for further information. MG11018401 Seq primer: custom primer used High quality sequence stop:182.

```
FEATURES      location/Qualifiers
source        1. .654
              /organism="Mus musculus
              /strain="C57BL"
              /db_xref="taxon:10090"
```

```

/clone="IMAGE:2503549"
/clone_lib="Sugano mouse kidney mk1a"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pME18S-FL3; Site:1; DraIII
(CACCATGTG); Site-2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTCGCCCTTTTTTTTTTTTTTTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor (TGTGGCCCTACTG), digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACCATGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Site selection was
performed to exclude the cDNA insert. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCCTCTCTTAAGACGCGC and 3' end
primer GCACCTGACGCTGACCAAC."
156 a 193 c 131 t
BASE COUNT
ORIGIN

```

[illegible]

Search completed: March 6, 2003, 19:37:28
Job time : 2189.5 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 15:27:01 ; Search time 65.5 Seconds
(without alignments)
7879.954 Million cell updates/sec

Title: US-09-912-559-1

Perfect score: 1683

Sequence: 1 atgttgcagagatgtctga.....aaagtgaagtgcttctaa 1683

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued_Patents_NA:*

2: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*

3: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*

5: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*

6: /cgn2_6/ptodata/1/ina/PCtus.COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	ID	Description
1	117.6	7.0	2033	1	US-08-148-910-14
2	117.6	7.0	2033	1	US-08-148-937A-14
3	96.2	5.7	970	1	US-08-148-910-3
4	96.2	5.7	970	1	US-08-148-937A-3
5	75.4	4.5	1065	1	US-08-427-640-1
6	75.4	4.5	1065	1	US-08-427-640-5
7	75.4	4.5	1068	1	US-08-427-640-3
8	75.4	4.5	1137	4	US-09-553-498-9
9	75.4	4.5	1137	4	US-09-618-869-9
10	75.4	4.5	1314	2	US-08-811-949-48
11	75.4	4.5	1955	2	US-08-883-795A-39
12	75.4	4.5	2457	6	5344773-1
13	75.4	4.5	7360	1	US-08-286-740-1
14	75.4	4.5	7360	1	PCT-US95-09576-1
15	73.8	4.4	329	1	US-08-148-910-2
16	73.8	4.4	329	1	US-08-148-910-13
17	73.8	4.4	329	1	US-08-148-910-13
18	73.8	4.4	329	1	US-08-448-937A-2
19	73.8	4.4	1065	2	US-08-448-937A-13
20	73.8	4.4	1068	1	US-08-811-949-60
21	73.8	4.4	1068	1	US-08-137-116-2
22	73.8	4.4	1068	2	US-08-427-640-7
23	73.8	4.4	1068	2	US-08-811-949-44
24	73.8	4.4	1068	2	US-08-811-949-46
25	73.8	4.4	1068	2	US-08-811-949-52
26	73.8	4.4	1068	2	US-08-811-949-58
27	73.8	4.4	1068	6	523256-3
					Patent No. 523256
					Sequence 5, Appli

28	73.8	4.4	1163	4	US-09-410-882-5	Sequence 5, Appli
29	73.8	4.4	1170	2	US-08-811-949-64	Sequence 64, Appl
30	73.8	4.4	1170	2	US-08-811-949-66	Sequence 66, Appl
31	73.8	4.4	1114	2	US-08-811-949-50	Sequence 50, Appl
32	73.8	4.4	1114	2	US-08-811-949-54	Sequence 54, Appl
33	73.8	4.4	1114	2	US-08-811-949-56	Sequence 56, Appl
34	73.8	4.4	1114	2	US-08-811-949-62	Sequence 62, Appl
35	73.8	4.4	1738	6	5200340-1	Patent No. 5200340
36	73.8	4.4	1848	3	US-08-814-412-10	Sequence 10, Appl
37	73.8	4.4	1974	2	US-08-811-949-38	Sequence 38, Appl
38	73.8	4.4	2101	2	US-08-811-949-42	Sequence 42, Appl
39	73.8	4.4	2162	1	US-08-119-512-3	Sequence 3, Appli
40	73.8	4.4	2162	1	US-08-488-0158-3	Sequence 54, Appl
41	73.8	4.4	2162	1	US-08-488-0158-25	Sequence 25, Appl
42	67.8	4.0	1724	6	5200340-5	Patent No. 5200340
43	67.8	4.0	2497	6	5185259-2	Patent No. 5185259
44	66.4	3.9	2544	4	US-09-518-046-3	Sequence 3, Appli
45	66	3.9	1233	1	US-08-254-922-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-148-910-14
Sequence 14, Application US/08148910
Patent No. 5465593
GENERAL INFORMATION:
APPLICANT: Takeshi SHIMOMURA et al.
TITLE OF INVENTION: No. 5465593el Protein and Gene Encoding Said Protein
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch,
MEDIUM TYPE: 500 Kb Storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/148,910
FILING DATE: No. 5465593ember 5, 1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX: 202-371-8856
TELEX:
INFORMATION FOR SEQ ID NO. 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 2033 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: linear
TOPOLOGY: linear
MOLECULE TYPE: CDNA
ORIGINAL SOURCE: human
ORGANISM: human
IMMEDIATE SOURCE:
LIBRARY: Pre-made Lambda phage library,
LIBRARY: human liver(49, male) CDNA Library (Stratagene)
US-08-148-910-14

Query Match	Similarity	7.0%	Score 117.6	DB 1	Length 2033
Best Local	Similarity	49.8%	Pred. 6.8e-25		
Matches	607	Conservative	0	Mismatches	559
				Indels	54
				Gaps	10
QY	469	AACCCCTGCCAAGATGGGGCTACCTGCTCCCGCATAGAGCGAATCCAAAGTCAACCTGT	528		
Db	742	AGCCCTGTGTGAAAGGGGGCACCTTCCACCTGATCTGGGCCACCGGACACCGTGTGT	801		
QY	529	GGCTGCCGACCAAGTTCACAGGGGAATTCGTGAAATPAGTTTCATCA---CTGGCAT	585		
Db	802	GCTGCCACCAAGGCTTCTGCTGAGCGGCTCTGTCAACATCAGCCTTATAGCGCTTCT	861		
QY	586	GTTCGCGATGGCTACTCTTACGAGGGGAAATGAATAGCAAGTCAACCGCATGCGTGC	645		
Db	862	TTGGGGAAGCGACTGGGTACCGTGGCGCGGACGACACACACCTCGGGCTCAAGCTGC	921		
QY	646	CTTTACTGGAACCTCCACCTCTCTTTCGAGGAAATTACAAATGTGTTAGGAGATGCT	705		
Db	922	CTGGCCTGGAACTCCGATCTGCTTACAGAGAGCTGCACGCTGGATCTCGTGGCGCGCG	981		
QY	706	GAATCCATGGAGATTGGGGAACAAATTTCTGCAAAAACCCAGATCGGACGAAAACCC	765		
Db	982	GGCCTCTGGGCGCTGGGCCCCCATGGCTCTGTCGCGAATCCGGACATATACAGAGGCTC	1041		
QY	766	TGGTGTATTATTAAGTTACCATGTACAAAGGAAATGGGAATACGTGATGTCTCAGCC	825		
Db	1042	TGGTGTAGCT---GGTGAAGAACACCGCGCTCTCTGGGAGTATCTCGCGCTTGAAGGC	1098		
QY	826	TGCTCAGCCCAAGACGTTTGCCTTACCAGAGAAACCCCACTAGCATCAACCAAGCTT	885		
Db	1099	TGGGAATCCCT-----CACCAAGTCCAACTGTCACCGGATCTCTTGGCGACCTGT	1149		
QY	886	CCGGGGTTGACTCTGTGGAAGAAGTGAATAGACAGAGAAATGTCAGAGAATCTAT	945		
Db	1150	CTGAGACAGCCTCCCGGGGACCAGGCGCTGGCAGAGAGGACAAAGAAGAGCGTTC	1209		
QY	946	GAGAGCTTTAAGACAGCGGGGCGCAAGCAACCATGGACGCTCCCTCAGTCTCGGTG	1005		
Db	1210	CTCGGCGCACGATATATCGGGGCTCTCTCTGCTGCGCGGCTGGACCCCTG-----	1262		
QY	1006	CCCTTGACCATCTCCATGCCCCAGGGGCCATTTGTGTGGGGCGCTGATCCACCCCTGC	1065		
Db	1263	--GCTGGCGGCATCTACATCGGGGACACTCTGCGCGGAGCGCTGGTCCACACTGC	1320		
QY	1066	TGGGTCTCAGCTGCTGGGCCACTGACCGACATATAAAACCAAGACATCTAA-----	1119		
Db	1321	TGGGTGTGTCTGGCGGCCACTGCTTCTTCCACAGCCCCCAGAGGACAGGCTCTCGGTG	1380		
QY	1120	GTGCTAGGGGACAGGACCTGAAGAAAGAATTTGATGAGCAGAGCTTTAGGGTGGAG	1179		
Db	1381	GTGCTGGGCGACACTTCTTCAACCGCACAGCGACGCTGACGAGACCTTTCGGCATGGAG	1440		
QY	1180	AAGATTTTCAAGTACAGCCACTCAATGAAAGAGTGAATGATCCCAACATATATTGCA	1239		
Db	1441	AAGTATATCCGCTACACCTGTACTGTGGTGTTCACACCCACAGCACAC---GACTGTGTC	1497		
QY	1240	TTGCTCAGGTTAAAGCCAGTGTATGTCTACTGTGCTGTAGATTCAAATATCTGAAGCT	1299		
Db	1498	CTGATCCGGCTGAAGAAAGAGGGGACCGCTGTGTGCACACGCTCCGACATGTTCTGACGCC	1557		
QY	1300	GTGTGCTTCCCGCAT-----GGTGCCTTTCCTGTGGAGTGAATGTCCACATCTCGGGC	1353		
Db	1558	ATGTGCTGCCGAGCCCGGACAGCTTCCCGGAGBACAAAGTGTCCAGATTGTGGGGC	1617		
QY	1354	TGGGG-----TGTTCAGAAACAGGAAAGGATCCCGCCACAGCTCTGTGATGCCAAATGTC	1407		
Db	1618	TGGGGGCACTGTGATGAGAAAGTGAACGTAACGAGCTACTCCAGCTCTCTGGGAGGACCTTGATC	1677		
QY	1408	AAGCTGATGCGCAACACTTTGTGCALCTCCGCCCACACTATACACATGATGTATGATGAC	1467		
Db	1678	CCCTGTGTGCGGACCAAGTGTGACAGCCCTGAGGTCTTAGGGCCCGACATCAGGCCCC	1737		
QY	1468	AGTATGATCTGTGACAGAAATCTTTCAGAAAACCTGGGCGAAGACACTGCCAGGCTGATCTT	1527		

[illegible]

```

1 RESULT 2
2 US-08-448-937A-14
3 ; Sequence 14, Application US/08448937A
4 ; Patent No. 5677164
5 ; GENERAL INFORMATION:
6 ; APPLICANT: Takeshi SHIMOMURA et al.
7 ; TITLE OF INVENTION: No. 5677164el Protein and Gene Encoding Said Protein
8 ; NUMBER OF SEQUENCES: 14
9 ; CORRESPONDENCE ADDRESS:
10 ; ADDRESSEE: Wenderoth, Lind & Ponack
11 ; STREET: 805 Fifteenth Street, N.W., #700
12 ; CITY: Washington
13 ; STATE: D.C.
14 ; COUNTRY: U.S.A.
15 ; ZIP: 20005
16 ; COMPUTER READABLE FORM:
17 ; MEDIUM TYPE: Diskette, 5.25 inch,
18 ; MEDIUM TYPE: 500 Kb Storage
19 ; COMPUTER: IBM Compatible
20 ; OPERATING SYSTEM: MS-DOS
21 ; SOFTWARE: Wordperfect
22 ; CURRENT APPLICATION DATA:
23 ; APPLICATION NUMBER: US/08/448,937A
24 ; FILING DATE: May 24, 1995
25 ; CLASSIFICATION: 435
26 ; PRIOR APPLICATION DATA:
27 ; APPLICATION NUMBER: 08/148,910
28 ; FILING DATE: No. 5677164ember 5, 1993
29 ; ATTORNEY/AGENT INFORMATION:
30 ; NAME: Warren M. Cheek, Jr.
31 ; REGISTRATION NUMBER: 33,367
32 ; REFERENCE/DOCKET NUMBER:
33 ; TELECOMMUNICATION INFORMATION:
34 ; TELEPHONE: 202-371-8850
35 ; TELEFAX: 202-371-8856
36 ; TELEX:
37 ; INFORMATION FOR SEQ ID NO: 14:
38 ; SEQUENCE CHARACTERISTICS:
39 ; LENGTH: 2033 base pairs
40 ; TYPE: nucleic acid
41 ; STRANDEDNESS: double
42 ; TOPOLOGY: linear
43 ; MOLECULE TYPE: cDNA
44 ; ORIGINAL SOURCE:
45 ; ORGANISM: human
46 ; IMMEDIATE SOURCE:
47 ; LIBRARY: Pre-made Lambda phage Library,
48 ; LIBRARY: human liver(49, male) cDNA Library (Stratagene)
49 ; US-08-448-937A-14
50
51 Query Match 7.0%; Score 117.6; DB 1; Length 2033;
52 Best Local Similarity 49.8%; Pred. No. 6,8e-25;
53 Matches 607; Conservative 0; Mismatches 559; Indels 54; Gaps 10
54
55 469 AACCCCTGCCACAATGAGGGCTACTGCTCCCGGCAATAAGCGAGATCCAAAGTTCACACTGT 528
56 ++++++|||||+++++|||||+++++|||||+++++|||||+++++|||||+++++|||||
57 742 AGCCCTTGCTCTTAACGGGGGCGACCTGCGACCTGATCGTGGCGACCGGGGACACACCGCTGT 801

```



```
OY 1028 AGGGCCACTTCTGTGTGGGGGCTGATCCACCCCTGCTGGTGTCTACCTGCTGCCACT 1087
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 461 AGCGTTCCTGTGCGGGGCGATATCATCTACCTCTCTCTGATTTCTCTGCGGCCACT 520
OY 1088 GCACCGA-----CATAAAACCAACATCTAAAGGTGTGTAGGGGACCGAGCTCGA 1141
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 521 GCTTCCAGAGAGGTTCCGCCACACCTGACGTGATGATCTTGCGGACAGACATACCGG 580
OY 1142 AGAAGAAGATTTTCATGACAGCTTTAGGTGAGAGATATTCAAGTACAGCCACT 1201
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 581 TGTTCCTGGCGAGAGAGAGAGAAATTTGAAGTGAATAATCATTTGCCATTAAGAT 640
OY 1202 ACATGAAAGAGATGAGATTTCCCAATGATATTGCTTCAAGTTAAAGCCAGTGG 1261
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 641 TCGAT-----GATGACACTTACGACATGACATTTGGGTGCGACGCTGAATGCAATT 694
OY 1262 ATGGTCACTGTGCTTCAAGATCCAAATACGTAAGACTGTGTGCTGCTGATGGTCT 1321
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 695 CGTCCGCTGTGCCAGAGAGACGCTGTCCGACTGTGTCTTCCCGCGCGAGCC 754
OY 1322 TTCCCT-----CTGGAGTGTGAGTCCCATCTCTGCTGGGTGTTACAGAAA 1369
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 755 TGCAGCTGCGGAGCTGACGAGAGTGTGCTCTCCGGCTACGGCAAGCATGAGGCTTGT 814
OY 1370 CAGGAAAAGGTCGCCGACGCTCTGATGCCAAAGTCAAGCTGATTGCCAAGCTTGT 1429
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 815 CTCCTTCTATTGAGAGCGGCTGAAGGAGGCTCATGTACAGCTGTACCATCAGCCGCT 874
OY 1430 GCACACTCCGGCCCACTGTATGACCATGATTTGATGACATATGATCTGTGAGAAATC 1489
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 875 GCACATACACATTTACTTTACAGACAGATCCAGCATGCTGTGTGTGAGACA 934
OY 1490 TTCAAAACCTG-----GCAAGACACTCCAGGGTGTGCTGTGAGAGCC 1534
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 935 CTCGAGCGCGGCGGCCAGGCAACTTGCACGACGCTCCAGGCGCATGAGGAGGCC 994
OY 1535 CCTGACCTGTGAGAGAGCGACCTACTACTGCTATGAGGATGAGTGGGGCCCTGG 1594
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 995 CCGTGTGTCTGACAGATGTCGCCCATGACTTTGGTGGCATCATCAGCTGGGCCCTGG 1054
OY 1595 AGTGTGGGAAGAG-----CCAGGGGTCTTACACCCCAAGTTACCAATTTCTGAATTGA 1648
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1055 GCTGTGACGAAAGATGTCGGGTGTGTACACCAAGTTACCACTACCTAGACTGGA 1114
OY 1649 TCAAGCCACCAT 1661
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1115 TTCTGTACACAT 1127
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
US-09-618-869-9
; Sequence 9, Application US/09618869
; Patent No. 6455279
; GENERAL INFORMATION:
; APPLICANT: Ambrosius, Dorothee
; APPLICANT: Rudolph, Rainer
; APPLICANT: Schaeffner, Joerg
; APPLICANT: Schwarz, Elisabeth
; TITLE OF INVENTION: PROCESS FOR THE PRODUCTION OF NATURALLY FOLDED AND
; TITLE OF INVENTION: SECRETED PROTEINS BY CO-SECRETION OF MOLECULAR
; FILE REFERENCE: 20381
; CURRENT APPLICATION NUMBER: US/09/618, 869
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: EP99114811.5
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1137
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
```

```
NAME/KEY: CDS
LOCATION: (1)..(1137)
US-09-618-869-9

Query Match
Best Local Similarity 50.8%; Score 75.4; DB 4; Length 1137;
Matches 342; Conservative 0; Mismatches 286; Indels 45; Gaps 5;

OY 1028 AGGGCCACTTCTGTGTGGGGGCGGTGATCCACCCCTGCTGGTGTCTACCTGCTGCCACT 1087
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 461 AGCGTTCCTGTGCGGGGCGATATCATCTACCTCTCTCTGATTTCTCTGCGGCCACT 520
OY 1088 GCACCGA-----CATAAAACCAACATCTAAAGGTGTGTAGGGGACCGAGCTCGA 1141
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 521 GCTTCCAGAGAGGTTCCGCCACACCTGACGTGATGATCTTGCGGACAGACATACCGG 580
OY 1142 AGAAGAAGATTTTCATGACAGCTTTAGGTGAGAGATATTCAAGTACAGCCACT 1201
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 581 TGTTCCTGGCGAGAGAGAGAGAAATTTGAAGTGAATAATCATTTGCCATTAAGAT 640
OY 1202 ACATGAAAGAGATGAGATTTCCCAATGATATTGCTTCAAGTTAAAGCCAGTGG 1261
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 641 TCGAT-----GATGACACTTACGACATGACATTTGGGTGCGACGCTGAATGCAATT 694
OY 1262 ATGGTCACTGTGCTTCAAGATCCAAATACGTAAGACTGTGTGCTGCTGATGGTCT 1321
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 695 CGTCCGCTGTGCCAGAGAGACGCTGTCCGACTGTGTGCTTCCCGCGCGAGCC 754
OY 1322 TTCCCT-----CTGGAGTGTGAGTCCCATCTCTGCTGGGTGTTACAGAAA 1369
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 755 TGCAGCTGCGGAGCTGACGAGAGTGTGCTCTCCGGCTACGGCAAGCATGAGGCTTGT 814
OY 1370 CAGGAAAAGGTCGCCGACGCTCTGATGCCAAAGTCAAGCTGATTGCCAAGCTTGT 1429
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 815 CTCCTTCTATTGAGAGCGGCTGAAGGAGGCTCATGTACAGCTGTACCATCAGCCGCT 874
OY 1430 GCACACTCCGGCCCACTGTATGACCATGATTTGATGACATATGATCTGTGAGAAATC 1489
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 875 GCACATACACATTTACTTTACAGACAGATCCAGCATGCTGTGTGTGAGACA 934
OY 1490 TTCAAAACCTG-----GCAAGACACTCCAGGGTGTGCTGTGAGAGCC 1534
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 935 CTCGAGCGCGGCGGCCAGGCAACTTGCACGACGCTCCAGGCGCATGAGGAGGCC 994
OY 1535 CCTGACCTGTGAGAGAGCGACCTACTACTGCTATGAGGATGAGTGGGGCCCTGG 1594
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 995 CCGTGTGTCTGACAGATGTCGCCCATGACTTTGGTGGCATCATCAGCTGGGCCCTGG 1054
OY 1595 AGTGTGGGAAGAG-----CCAGGGGTCTTACACCCCAAGTTACCAATTTCTGAATTGA 1648
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1055 GCTGTGACGAAAGATGTCGGGTGTGTACACCAAGTTACCACTACCTAGACTGGA 1114
OY 1649 TCAAGCCACCAT 1661
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1115 TTCTGTACACAT 1127
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
US-08-811-949-48
; Sequence 48, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIMA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESS: P.C.
```

STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: OBOLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220

INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 1314 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)

FEATURE:
NAME/KEY: CDS
LOCATION: 1..1311
US-08-811-949-48


```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 7360 bases
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-286-740-1

```

```

Query Match          4.5%; Score 75.4; DB 1; Length 7360;
Best Local Similarity 50.8%; Pred. No. 4.4e-12;
Matches 342; Conservative 0; Mismatches 286; Indels 45; Gaps 5;

```

```

OY 1028 AGGGCCACTTCTGTGTGGGCGCTGATCCACCCCTGGGTGCTCACTGCTGCCACT 1087
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2634 AGGGTTCCTGTGGGGGCGCATATCATACGCTCGTGATTCCTGTGCGGCCCACT 2693
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1088 GCACCCA-----CATMAAACCGACATCTAAAGTGTGTAGGGGACCAAGCCTGA 1141
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2694 GCTTCCAGAGAGGTTTCCGCCCCACCTGACGTGATCTTGGGCGAACAATACCGGG 2753
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1142 AGAAGAAGATTTTCATGACGACAGCTTTAGGGTGAGAGATATTCAGTACAGCCACT 1201
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2754 TGGTCCCTGGCAGAGAGAGAGCAAAATTTGAAGTCGAAAATAATATGTCATAGGAAT 2813
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1202 ACAATGAAAGATAGATTTCCCAATATATTTGCAATGCTCAAGTTAAGCCACTGG 1261
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2814 TCGAT-----GATGACACTTACGACATGACATTCGCTGCGACGCTGAATCGAAT 2867
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1262 ATGATCACTGCTCTAGAAATCAAAATACGTGAAGACTGTGCTTGGCTGATGGTCT 1321
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2868 CGTCCGCTGCGCCAGAGAGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2927
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1332 TTGCTT-----CTGGGAGTAGTCCCACTCTCTGGCTGGGGGTGTTTCAGAAA 1369
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2928 TGCAGCTGCCGAGCTGAGACGAGTGTAGCTCTCGGCTGACGCAAGATAGGCTTGT 2987
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1370 CAGGAAAGGGTCCCGCCAGCTCTCTGATGCCAAATGCAAGTCAAGTCCCACTTTGT 1429
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2988 CTCCTTTTATTTCGAGCGGCTGAGAGAGGCTCATGTACACTGTAACCATCCAGCCGCT 3047
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1430 GCACATCCGCCCACTATATGACCAATGATGATGACATGATGATGATGATGATGATG 1489
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3048 GCACATCACAACATTTACTTAAACAAACAGTCAACCAACATGCTGTGTGTGGAACA 3107
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1490 TTCGAAACCTGG-----GCAAGACCTGCGCAGGGTCACTCTGGAAGGCC 1534
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3108 CTCGAGACGGGGGCCAGCCAGCAACTTTCAGCAGACGCTGCGCAGGATTCGGAGGGC 3167
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1535 CCCGACCTGTGGAAGAGCGCACTACTACGTATGATGATGATGATGATGATGATGATG 1594
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3168 CCTGTGTGTCTGAACGATGCGCCCATGACTTGTGGTGGGCAATCATGAGTGGGCTGG 3227
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1595 AGTGTGGGAAGG-----CCAGGGGTCTACACCCAAAGTTACCAATTTCTGAATTGA 1648
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3228 GCTGTGAGACAGAGATGTCCGGGTGTGTAACACAAAGTTACCAACTACTAGACTGA 3287
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1649 TCAGAGCCACCAT 1661
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3288 TTCGTGACACAT 3300
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 14
PCT-US95-09576-1
; Sequence 1, Application PC/TUS9509576
; GENERAL INFORMATION:
; APPLICANT: GENENTECH, INC.
; TITLE OF INVENTION: METHOD FOR SELECTING HIGH-EXPRESSING
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California

```

```

; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09576
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/286740
; FILING DATE: 05-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: 798PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; INFORMATION FOR SEQ. ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7360 bases
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
PCT-US95-09576-1

```

```

Query Match          4.5%; Score 75.4; DB 5; Length 7360;
Best Local Similarity 50.8%; Pred. No. 4.4e-12;
Matches 342; Conservative 0; Mismatches 286; Indels 45; Gaps 5;

```

```

OY 1028 AGGGCCACTTCTGTGTGGGCGCTGATCCACCCCTGGGTGCTCACTGCTGCCACT 1087
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2634 AGGGTTCCTGTGGGGGCGCATATCATACGCTCGTGATTCCTGTGCGGCCCACT 2693
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1088 GCACCCA-----CATMAAACCGACATCTAAAGTGTGTAGGGGACCAAGCCTGA 1141
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2694 GCTTCCAGAGAGGTTTCCGCCCCACCTGACGTGATCTTGGGCGAACAATACCGGG 2753
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1142 AGAAGAAGATTTTCATGACGACAGCTTTAGGGTGAGAGATATTCAGTACAGCCACT 1201
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2754 TGGTCCCTGGCAGAGAGAGAGCAAAATTTGAAGTCGAAAATAATATGTCATAGGAAT 2813
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1202 ACAATGAAAGATAGATTTCCCAATATATTTGCAATGCTCAAGTTAAGCCACTGG 1261
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2814 TCGAT-----GATGACACTTACGACATGACATTCGCTGCGACGCTGAATCGAAT 2867
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1262 ATGATCACTGCTCTAGAAATCAAAATACGTGAAGACTGTGCTTGGCTGATGGTCT 1321
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2868 CGTCCGCTGTCGCCAGAGAGAGAGGTGTGCGACTGTGTGCTTCCCGGGGGAGC 2927
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1332 TTGCTT-----CTGGGAGTAGTCCCACTCTCTGGCTGGGGTGTACAGAAA 1369
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2928 TGCAGCTCCCGACTGGACGAGTGTGACCTCTCGGCTAGCGCAACATGAGGCCCTGT 2987
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1370 CAGGAAAGGGTCCCGCCAGCTCTCTGATGCCAAATGCAAGTCAAGTCCCACTTTGT 1429
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2988 CTCCTTTTATTTCGAGCGGCTGAGAGAGGCTCATGTGATGATGATGATGATGATGATG 3047
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1430 GCACATCCGCCCACTATATGACCAATGATGATGATGATGATGATGATGATGATGATG 1489
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3048 GCACATCACAACATTTACTTAAACAAACAGTCAACCAACATGCTGTGTGGAACA 3107
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1490 TTCGAAACCTGG-----GCAAGACCTGCGCAGGGTCACTCTGGAAGGCC 1534
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3108 CTCGAGACGGGGGCCAGCCAGCAACTTTCAGCAGACGCTGCGCAGGATTCGGAGGGC 3167
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1535 CCCGACCTGTGGAAGAGCGCACTACTACGTATGATGATGATGATGATGATGATGATGATG 1594
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3168 CCTGTGTGTCTGAACGATGCGCCCATGACTTGTGGTGGGCAATCATGAGTGGGCTGG 3227
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 18:44:22 ; Search time 127.5 seconds
(without alignments)

8816.543 Million cell updates/sec

Title: US-09-912-559-1

Perfect score: 1683
Sequence: 1 atgtttccagagatgcttctga.....aaagtaaaagtgcttctaa 1683

Scoring table: IDENTITY_NUC
Gapop 10.0 , Capext 1.0

Searched: 478924 seqs, 333959956 residues

Total number of hits satisfying chosen parameters: 957848

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCIT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCITUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1683	100.0	1683	10	US-09-912-559-1	Sequence 1, Appl1
2	1683	100.0	3008	10	US-09-880-107-1668	Sequence 1668, Ap
3	1679.8	99.8	1683	10	US-09-912-559-2	Sequence 2, Appl1
4	120.2	7.1	451	10	US-09-864-761-11164	Sequence 1164, A
5	117.6	7.0	2036	10	US-09-954-456-552	Sequence 552, Ap
6	117.6	7.0	2036	10	US-09-880-107-1612	Sequence 1612, Ap
7	117	7.0	1117	10	US-09-864-761-27791	Sequence 27791, A
8	75.4	4.5	1689	10	US-09-969-271-6	Sequence 6, Appl1
9	75.4	4.5	2519	10	US-09-969-271-5	Sequence 5, Appl1
10	75.4	4.5	2519	10	US-09-974-298-144	Sequence 144, Ap
11	72.8	4.3	614	10	US-09-879-792-33	Sequence 33, Appl
12	68	4.0	1614	10	US-09-888-615-45	Sequence 45, Appl
13	68	4.0	1748	10	US-09-879-792-11	Sequence 11, Appl
14	66.4	3.9	1230	10	US-09-879-792-35	Sequence 35, Appl
15	66.4	3.9	2412	9	US-09-878-295A-68	Sequence 68, Appl
16	66.4	3.9	2412	9	US-09-878-697-68	Sequence 68, Appl
17	66.4	3.9	2412	9	US-09-878-192A-68	Sequence 68, Appl
18	66.4	3.9	2412	9	US-09-999-832A-68	Sequence 68, Appl
19	66.4	3.9	2412	9	US-09-978-189-68	Sequence 68, Appl

20	66.4	3.9	2412	9	US-10-174-590-63	Sequence 63, Appl
21	66.4	3.9	2412	9	US-10-176-758-63	Sequence 63, Appl
22	66.4	3.9	2412	9	US-10-175-737-63	Sequence 63, Appl
23	66.4	3.9	2412	9	US-10-173-706-63	Sequence 63, Appl
24	66.4	3.9	2412	9	US-10-175-738-63	Sequence 63, Appl
25	66.4	3.9	2412	9	US-10-175-752-63	Sequence 63, Appl
26	66.4	3.9	2412	9	US-10-176-482-63	Sequence 63, Appl
27	66.4	3.9	2412	9	US-10-176-757-63	Sequence 63, Appl
28	66.4	3.9	2412	9	US-10-176-913-63	Sequence 63, Appl
29	66.4	3.9	2412	9	US-10-180-552-63	Sequence 63, Appl
30	66.4	3.9	2412	9	US-10-180-557-63	Sequence 63, Appl
31	66.4	3.9	2412	9	US-10-173-700-63	Sequence 63, Appl
32	66.4	3.9	2412	9	US-10-174-572-63	Sequence 63, Appl
33	66.4	3.9	2412	9	US-10-174-579-63	Sequence 63, Appl
34	66.4	3.9	2412	9	US-10-174-582-63	Sequence 63, Appl
35	66.4	3.9	2412	9	US-10-174-588-63	Sequence 63, Appl
36	66.4	3.9	2412	9	US-10-175-739-63	Sequence 63, Appl
37	66.4	3.9	2412	9	US-10-175-740-63	Sequence 63, Appl
38	66.4	3.9	2412	9	US-10-175-743-63	Sequence 63, Appl
39	66.4	3.9	2412	9	US-10-176-488-63	Sequence 63, Appl
40	66.4	3.9	2412	9	US-10-176-492-63	Sequence 63, Appl
41	66.4	3.9	2412	9	US-10-176-747-63	Sequence 63, Appl
42	66.4	3.9	2412	9	US-10-176-750-63	Sequence 63, Appl
43	66.4	3.9	2412	9	US-10-176-985-63	Sequence 63, Appl
44	66.4	3.9	2412	9	US-10-176-987-63	Sequence 63, Appl
45	66.4	3.9	2412	9	US-10-176-991-63	Sequence 63, Appl

ALIGNMENTS

RESULT 1
US-09-912-559-1
Sequence 1, Application US/09912559
Patent No. US20020142316A1
GENERAL INFORMATION:
APPLICANT: ROEMISCH, JOERGEN
APPLICANT: STOEHR, HANS-ARNOLD
APPLICANT: FEUSSNER, ANNETTE
APPLICANT: LANG, WIEGAND
APPLICANT: WEIMER, THOMAS
APPLICANT: BECKER, MARGRET
APPLICANT: NEPLICH, CLAUDIA
APPLICANT: MUTH-NAUMANN, GUDRUN
TITLE OF INVENTION: MUTANTS OF THE FACTOR VII-ACTIVATING PROTEASE AND
FILE OF INVENTION: DETECTION METHODS USING SPECIFIC ANTIBODIES
FILE REFERENCE: 06478.1457
CURRENT APPLICATION NUMBER: US/09/912,559
CURRENT FILING DATE: 2001-07-26
PRIOR APPLICATION NUMBER: DE 100 36 641.4
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: DE 100 50 040.4
PRIOR FILING DATE: 2000-10-10
PRIOR APPLICATION NUMBER: DE 100 52 319.6
PRIOR FILING DATE: 2000-10-21
PRIOR APPLICATION NUMBER: DE 101 18 706.8
PRIOR FILING DATE: 2001-04-12
NUMBER OF SEQ. ID NOS: 4
SOFTWARE: PatentIn Ver. 2.1
SEQ. ID NO. 1
LENGTH: 1683
TYPE: DNA
ORGANISM: Homo sapiens
US-09-912-559-1

Query Match 100.0%; Score 1683; DB 10; Length 1683;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1683; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTTCCAGAGATGCTGATCTCCATGTTGCTGTTAAATGCTGTGGTGAAGACA 60
DB 1 ATGTTTCCAGAGATGCTGATCTCCATGTTGCTGTTAAATGCTGTGGTGAAGACA 60

QY	61	GCCTGGGGTTCCTCCATCTCTTTATTTGGAAGCTGGACCCAGACTGGACCCCTGAC	120
Db	61	GCCTGGGGTTCCTCCATCTCTTTATTTGGAAGCTGGACCCAGACTGGACCCCTGAC	120
QY	121	CAGTATGATTTACAGTACGAGAGATTATTAATCAGAGAAAGAACACCAGTACACTTACC	180
Db	121	CAGTATGATTTACAGTACGAGAGATTATTAATCAGAGAAAGAACACCAGTACACTTACC	180
QY	181	CATGCTGAGAAATCTGACTGGTACTACACTGAGAGACCAAGCTGATCCAGGCCAAC	240
Db	181	CATGCTGAGAAATCTGACTGGTACTACACTGAGAGACCAAGCTGATCCAGGCCAAC	240
QY	241	CCCTGTGAACACGGTGGGAGATGCTCTCTGTCATGGGAGCACCCTTCATATGACTGCTG	300
Db	241	CCCTGTGAACACGGTGGGAGATGCTCTCTGTCATGGGAGCACCCTTCATATGACTGCTG	300
QY	301	GCCTCTTTCTGTGGGAATTAAGTGTCAAGAAAGTCAAAAATACGTCGAAGACAACCCATGT	360
Db	301	GCCTCTTTCTGTGGGAATTAAGTGTCAAGAAAGTCAAAAATACGTCGAAGACAACCCATGT	360
QY	361	GGCGGGGGCCAAATGTCTCATATACCGAGAGGCTCCCTACTACCCCTGTGTCTTAACAC	420
Db	361	GGCGGGGGCCAAATGTCTCATATATCCCAAGAGCTCCCTCTACTACCCCTGTGTCTTAACAC	420
QY	421	CCTTACACAGATCCAGCTGCTCCCAAGATGATGTCTGTATCAGAGCCCAAAACCCCTGACG	480
Db	421	CCTTACACAGATCCAGCTGCTCCCAAGATGATGTCTGTATCAGAGCCCAAAACCCCTGACG	480
QY	481	AATGGGGCTACCTGCTGCCCGCAATAGCCGGAAGTCCAAAGTTACCTGTGCTGTCCGAC	540
Db	481	AATGGGGCTACCTGCTGCCCGCAATAGCCGGAAGTCCAAAGTTACCTGTGCTGTCCGAC	540
QY	541	CAGTTCAAGGGAAATTTGTGAATTAAGTTTGTGATGACTGCTATGTTGGCGATGGCTAC	600
Db	541	CAGTTCAAGGGAAATTTGTGAATTAAGTTTGTGATGACTGCTATGTTGGCGATGGCTAC	600
QY	601	TCTTACCGAGGGAAATTAATATAGACACGTCACACACAGATGGTGGCTTTACTGGAACCTCC	660
Db	601	TCTTACCGAGGGAAATTAATATAGACACGTCACACACAGATGGTGGCTTTACTGGAACCTCC	660
QY	661	CACCTCTCTTTCAGAGAGATTTACACATCTTTATGAGAGATGCTGTGAACCCATGGGATT	720
Db	661	CACCTCTCTTTCAGAGAGATTTACACATCTTTATGAGAGATGCTGTGAACCCATGGGATT	720
QY	721	GGGGAACACAAATTTGTGAGAAACCCAGATGGGAGAGAAAGCCCTGTGCTTTATTAA	780
Db	721	GGGGAACACAAATTTGTGAGAAACCCAGATGGGAGAGAAAGCCCTGTGCTTTATTAA	780
QY	781	GTTTACCAATGACAMGTGAATAATGGAAATACTGTGATGTCTCAGCCTTCAGGCCAGAC	840
Db	781	GTTTACCAATGACAMGTGAATAATGGAAATACTGTGATGTCTCAGCCTTCAGGCCAGAC	840
QY	841	GTTTGCTTACCCAGAGAAAGCCCACTGAGGCATCAACCAAGCTTCGCGGGTTTGACTCC	900
Db	841	GTTTGCTTACCCAGAGAAAGCCCACTGAGGCATCAACCAAGCTTCGCGGGTTTGACTCC	900
QY	901	TGTGGAAGACTGAGATGACAGAGAGAAAGATCAAGAGAATCTATGAGGGCTTTAAGAC	960
Db	901	TGTGGAAGACTGAGATGACAGAGAGAAAGATCAAGAGAATCTATGAGGGCTTTAAGAC	960
QY	961	ACGGCGGACAAACACCATGGACAGGGCTCCCTCAGTCCCTGCTGCTGACCAATCTCC	1020
Db	961	ACGGCGGACAAACACCATGGACAGGGCTCCCTCAGTCCCTGCTGCTGACCAATCTCC	1020
QY	1021	ATGCCCCAGGGCACCTTCTGTGTGTGGGGCGTGTATCCACCCCTGTGGGTCTCACTGCT	1080
Db	1021	ATGCCCCAGGGCACCTTCTGTGTGTGGGGCGTGTATCCACCCCTGTGGGTCTCACTGCT	1080
QY	1081	GGCCACTGCACGACATPAATAAACAACAGCACTPAAGGTGTGTGTAGGGGACCGAGACCTG	1140
Db	1081	GGCCACTGCACGACATPAATAAACAACAGCACTPAAGGTGTGTGTAGGGGACCGAGACCTG	1140
QY	1141	AMGAAAGAAATTTCATGACACAGCTTTAGGTGGAAGAAATATTCAGATACAGCCAC	1200

Db	1141	AGGAAAGAAATTTCTATGAGCAGACTTTAGGGTGGAGAAAGATTCAAGTACAGCCAC	1200
Qy	1201	TACATGAAAGAGATGAGATTTCCCAACAATGATATGGATTGCTCAAGTTAAAGCCACTG	1260
Db	1201	TACAAATGAAAGATGAGATTTCCCAACAATGATATGATTTGCTCAAGTTAAAGCCACTG	1260
Qy	1261	GATGGTCACTGTGCTTGAATCCAATACGTGAAGACGTGTGCTTGCTGCTGATGGGTCC	1320
Db	1261	GATGGTCACTGTGCTTGAATCCAATACGTGAAGACGTGTGCTTGCTGCTGATGGGTCC	1320
Qy	1321	TTTCCCTCTGGGAGTGAAGTGCACATCTGTGGCTGGGGTGTTCACAAACACAGAAAGAGG	1380
Db	1321	TTTCCCTCTGGGAGTGAAGTGCACATCTGTGGCTGGGGTGTTCACAAACACAGAAAGAGG	1380
Qy	1381	TCCCGCCACACTCTCGGATGCCAAGTCAAGCTGATTGCCAACACTTTGTGCACTCCCGC	1440
Db	1381	TCCCGCCACACTCTCGGATGCCAAGTCAAGCTGATTGCCAACACTTTGTGCAACTCCCGC	1440
Qy	1441	CAACTCTATGACCAATGATGATGACAGTATGATCTGTCCAGAGAAATCTTTCAGAAACT	1500
Db	1441	CAACTCTATGACCAATGATGATGACAGTATGATCTGTCCAGAGAAATCTTTCAGAAACT	1500
Qy	1501	GGGCAAGACACTCGCAGGGGTACTCTGGAAGGCCCTGACCTGTGAAGAGACGGCACCC	1560
Db	1501	GGGCAAGACACTCGCAGGGGTACTCTGGAAGGCCCTGACCTGTGAAGAGACGGCACCC	1560
Qy	1561	TACTAGCTATGSGGATAGTGAAGCTGGGGCTCGAGTGTGGGAAGAGGGCCAGGGGTCTAC	1620
Db	1561	TACTAGCTATGSGGATAGTGAAGCTGGGGCTCGAGTGTGGGAAGAGGGCCAGGGGTCTAC	1620
Qy	1621	ACCCAAAGTTACCAATTCCTGGAATTTGATCAAGGCACACATCAAAAGTCAAAAGTGGCTTC	1680
Db	1621	ACCCAAAGTTACCAATTCCTGGAATTTGATCAAGGCACACATCAAAAGTCAAAAGTGGCTTC	1680
Qy	1681	TAA 1683	
Db	1681	TAA 1683	

```

RESULT 2
US-09-880-107-1668
? Sequence 1668, Application US/09880107
? Patent No. US20020142981A1
? GENERAL INFORMATION:
? APPLICANT: Horne, Darci T.
? APPLICANT: Vockley, Joseph G.
? APPLICANT: Scherf, Uwe
? APPLICANT: Gene Logic, Inc.
? TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
? FILE REFERENCE: 44921-5028-WO
? CURRENT APPLICATION NUMBER: US/09/880,107
? CURRENT FILING DATE: 2001-06-14
? PRIOR APPLICATION NUMBER: US 60/211,379
? PRIOR FILING DATE: 2000-06-14
? PRIOR APPLICATION NUMBER: US 60/237,054
? PRIOR FILING DATE: 2000-10-02
? NUMBER OF SEQ ID NOS: 3950
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 1668
? LENGTH: 3008
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? OTHER INFORMATION: Genbank Accession No. US20020142981A1 D49742
US-09-880-107-1668

Query Match          100.0%; Score 1683; DB 10; Length 3008;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1683; Conservative 0; Mismatches 0; Indels 0; Gaps
0

1 ATGCTTGCCAGGATGTCGTGATCCATGCTTCGCTTTAATGCTCTGTGAGGAAAGACA 60
|||||

```

Db 97 ATGTTGCCAGATGCTGATCTCCATGTTCTGCTTAATGGCTCTGTGGAAAGACA 156
QY 61 GCCTGTGGGTCCCTGATGCTTTATTTGAAAGCTGAGCCAGACTGAGACCCCTGAC 120
Db 157 GCCTGTGGGTCTCCCTGATGCTTTATTTGAAAGCTGAGCCAGACTGAGACCCCTGAC 216
QY 121 CAGTATGATTACAGCTACGAGAGATTATTAATCAGAAAGAACACCAATGACACTTACC 180
Db 217 CAGTATGATTACAGCTACGAGAGATTATTAATCAGAAAGAACACCAATGACACTTACC 276
QY 181 CATGCTGAGAAATCTGACTGCTACTACACTGAGACCAAGTCAATGCGACACCAAC 240
Db 277 CATGCTGAGAAATCTGACTGCTACTACACTGAGACCAAGTCAATGCGACACCAAC 336
QY 241 CCCGTGTAACAGCGTGGGAGCTGCTGCTCATGGAGACACTTACATGAGCTGCTG 300
Db 337 CCCTGTGAACAGCGTGGGAGCTGCTGCTCATGGAGACACTTACATGAGCTGCTG 396
QY 301 GCTCCTTTCTCTGGAAATTAAGTGTACAGAAAGTCAAAATACGTGCAGACACCAATGT 360
Db 397 GCTCCTTTCTCTGGAAATTAAGTGTACAGAAAGTCAAAATACGTGCAGACACCAATGT 456
QY 361 GCGCGGGGCCAATGCTCATTAACAGAGTCCCTGCTACTACCGGTGTGTGTAAACAC 420
Db 457 GCGCGGGGCCAATGCTCATTAACAGAGTCCCTGCTACTACCGGTGTGTGTAAACAC 516
QY 421 CCTTACACAGGTCCCACTGCTCCCAAGTGTCTGTATGCAAGGCCAAACCCCTGCGAC 480
Db 517 CCTTACACAGGTCCCACTGCTCCCAAGTGTCTGTATGCAAGGCCAAACCCCTGCGAC 576
QY 481 AATGGGGCTACTGCTCCCGGATTAAGCGAGATCCAAAGTTCACCTGTGCTGCGGAC 540
Db 577 AATGGGGCTACTGCTCCCGGATTAAGCGAGATCCAAAGTTCACCTGTGCTGCGGAC 636
QY 541 CAGTTCAAGGGGAAATTTGTGAAATAGTTGATGATGCTGATGTGGGAGGTGTAC 600
Db 637 CAGTTCAAGGGGAAATTTGTGAAATAGTTGATGATGCTGATGTGGGAGGTGTAC 696
QY 601 TCTTACGAGGAAATGAAATAGCAAGTCAACCAATGCTGTGCTTTACTGTGAACTCC 660
Db 697 TCTTACGAGGAAATGAAATAGCAAGTCAACCAATGCTGTGCTTTACTGTGAACTCC 756
QY 661 CACCTCCTCTTGAGAGAAATTAACACATGTTTATGAGAGATGCTGAACCACTGGGATT 720
Db 757 CACCTCCTCTTGAGAGAAATTAACACATGTTTATGAGAGATGCTGAACCACTGGGATT 816
QY 721 GGGGAACAATTTCTGCAGAACCCAGATGCGAGCAAAAGCCCTGTGCTTTATTA 780
Db 817 GGGGAACAATTTCTGCAGAACCCAGATGCGAGCAAAAGCCCTGTGCTTTATTA 876
QY 781 GTTACCAATGACAAGGTGAATGGAATACTGATGCTCAAGCTGCTAGCCAGAC 840
Db 877 GTTACCAATGACAAGGTGAATGGAATACTGATGCTCAAGCTGCTAGCCAGAC 936
QY 841 GTTGGCTAACCCAGAAAGCCCACTGAGCCATCAACCAAGCTTCCGGGTGTGACTCC 900
Db 937 GTTGGCTAACCCAGAAAGCCCACTGAGCCATCAACCAAGCTTCCGGGTGTGACTCC 996
QY 901 TGTGGAAGAAGCTGAGTACAGAGAGAAATCAAGAAATCTATGAGAGCTTTAAGAGC 960
Db 997 TGTGGAAGAAGCTGAGTACAGAGAGAAATCAAGAAATCTATGAGAGCTTTAAGAGC 1056
QY 961 ACGGCGGGAAGCAACCATGGAGAGCTCCCTCAGCTCCCTGCTCTGACCATCTCC 1020
Db 1057 ACGGCGGGAAGCAACCATGGAGAGCTCCCTCAGCTCCCTGCTCTGACCATCTCC 1116
QY 1021 ATGCCCCAGGGCCACTTCTGTGTGGGGCGTGTATCCACCCTGCTGGGTCTTACTGCT 1080
Db 1117 ATGCCCCAGGGCCACTTCTGTGTGGGGCGTGTATCCACCCTGCTGGGTCTTACTGCT 1176
QY 1081 GCCCACTGACGACATTAATAAACAGACATCTAAAGGTGCTAGGGGACACAGACTG 1140
Db 1177 GCCCACTGACGACATTAATAAACAGACATCTAAAGGTGCTAGGGGACACAGACTG 1236

QY 1141 AAGAAGCAAGATTTTCATGAGCAGAGCTTTAGGTGGAGACAATTTCAAGTACAGCCAC 1200
Db 1237 AAGAAGCAAGATTTTCATGAGCAGAGCTTTAGGTGGAGACAATTTCAAGTACAGCCAC 1296
QY 1201 TACAATGAAGAGATGATTTCCCAATGATATTTGATTCATTCAGTTAAAGCCAGTG 1260
Db 1297 TACAATGAAGAGATGATTTCCCAATGATATTTGATTCATTCAGTTAAAGCCAGTG 1356
QY 1261 GATGTCACATGCTCTAGAAATCCAAATACGTGAAGACATGTGCTGCTGATGAGGTCC 1320
Db 1357 GATGTCACATGCTCTAGAAATCCAAATACGTGAAGACATGTGCTGCTGATGAGGTCC 1416
QY 1321 TTTCCTCTGGAGTACGAGCCACATCTGCTGCTGGGTGTACGAAGAACAGAAAGGG 1380
Db 1417 TTTCCTCTGGAGTACGAGCCACATCTGCTGCTGGGTGTACGAAGAACAGAAAGGG 1476
QY 1381 TCCCGCAGCTCCTGGATGCCAAAGTCAAGCTGATTGCCACACTTGTGCACATCCCGC 1440
Db 1477 TCCCGCAGCTCCTGGATGCCAAAGTCAAGCTGATTGCCACACTTGTGCACATCCCGC 1536
QY 1441 CAACCTATGACCATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
Db 1537 CAACCTATGACCATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1596
QY 1501 GGGCAAGACCTGCGAGGTGACTCTGAGAGGCCCTGACCTGTGAGAGAGACGGCAC 1560
Db 1597 GGGCAAGACCTGCGAGGTGACTCTGAGAGGCCCTGACCTGTGAGAGAGACGGCAC 1656
QY 1561 TACTACGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
Db 1657 TACTACGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1716
QY 1621 ACCCAAGTTACCAATTTCTGTAATGATCAAGCCACATCAAAAGTGAAGTGTTC 1680
Db 1717 ACCCAAGTTACCAATTTCTGTAATGATCAAGCCACATCAAAAGTGAAGTGTTC 1776
QY 1681 TAA 1683
Db 1777 TAA 1779

RESULT 3
US-09-912-559-2
Sequence 2, Application US/0912559
Patent No. US20020142316A1
GENERAL INFORMATION:
APPLICANT: ROEMISCH, JUERGEN
APPLICANT: STOEHR, HANS-ARNOLD
APPLICANT: FEUSSNER, ANNETTE
APPLICANT: LANG, WIEGAND
APPLICANT: WEIMER, THOMAS
APPLICANT: BECKER, MARGRET
APPLICANT: NEBLICH, CLAUDIA
APPLICANT: KUTH-NAUMANN, GUDRUN
TITLE OF INVENTION: MUTANTS OF THE FACTOR VII-ACTIVATING PROTEASE AND
TITLE OF INVENTION: DETECTION METHODS USING SPECIFIC ANTIBODIES
FILE REFERENCE: 06478.1457
CURRENT APPLICATION NUMBER: US/09/912.559
CURRENT FILING DATE: 2001-07-26
PRIOR APPLICATION NUMBER: DE 100 36 641.4
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: DE 100 50 040.4
PRIOR FILING DATE: 2000-10-10
PRIOR APPLICATION NUMBER: DE 100 52 319.6
PRIOR FILING DATE: 2000-10-21
PRIOR APPLICATION NUMBER: DE 101 18 706.8
PRIOR FILING DATE: 2001-04-12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 1683
TYPE: DNA

ORGANISM: Homo sapiens
US-09-912-559-2

Query Match 99.8%; Score 1679.8; DB 10; Length 1683;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1681; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGTTTCCAGAGATCTGATCTCATGTCTGCTTTAATGCTCTGTGGGAAAGACA 60
1 ATGTTGACAGATCTGATCTCATGTCTGCTTTAATGCTCTGTGGGAAAGACA 60
61 GCCTGGGCTTCCCTGATGCTTTATGGAAGCCTGGACCCAGATGACCCCTGAC 120
61 GCCTGGGCTTCCCTGATGCTTTATGGAAGCCTGGACCCAGATGACCCCTGAC 120
121 CAGTATGATTACAGTACGAGAGATTAATCAGGAAGACACAGATGACACTTACC 180
121 CAGTATGATTACAGTACGAGAGATTAATCAGGAAGACACAGATGACACTTACC 180
181 CATGCTGAGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 240
181 CATGCTGAGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 240
241 CCCTGTGAAACAGGTGGGAGCTCCCTGCTCATGGAGACCTTCACTGACGTGCTG 300
241 CCCTGTGAAACAGGTGGGAGCTCCCTGCTCATGGAGACCTTCACTGACGTGCTG 300
301 GCTCTTTCTCTGGGAATTAAGTGTGAGAAAGTGCATAAATACGTGCAAGGACCA 360
301 GCTCTTTCTCTGGGAATTAAGTGTGAGAAAGTGCATAAATACGTGCAAGGACCA 360
361 GGCCGGGGCCATGCTCTCTAATACCAAGATCTCTCTACTACCGCTGTGTCTGTAAC 420
361 GGCCGGGGCCATGCTCTCTAATACCAAGATCTCTCTACTACCGCTGTGTCTGTAAC 420
421 CCTTACACAGTCCAGCTGCTCCCAAGTGGTCTCTATGAGGCGCAACCCCTGAC 480
421 CCTTACACAGTCCAGCTGCTCCCAAGTGGTCTCTATGAGGCGCAACCCCTGAC 480
481 AATGGGCTCTCTGCTCCCGCATTAAGCGGAGATCCCAAGTTCACCTGTGCTCCGAC 540
481 AATGGGCTCTCTGCTCCCGCATTAAGCGGAGATCCCAAGTTCACCTGTGCTCCGAC 540
541 CAGTTCAAGGGGAATTTCTGTAATAGTCTGTAGTACTGTATGAGGAGTCTGAACCCAT 600
541 CAGTTCAAGGGGAATTTCTGTAATAGTCTGTAGTACTGTATGAGGAGTCTGAACCCAT 600
601 TCTTACGAGGAAATGATAGACACTCAACGATGCTGCTTACTGGAACCTCC 660
601 TCTTACGAGGAAATGATAGACACTCAACGATGCTGCTTACTGGAACCTCC 660
661 CACCTCCTCTGAGAGAAATTACAACATGTTTATGAGAGATCTGAACCCATGGGATT 720
661 CACCTCCTCTGAGAGAAATTACAACATGTTTATGAGAGATCTGAACCCATGGGATT 720
721 GGGGAACCAATTTCTGCAAAACCCAGATGGGAGCAAAAGCCCTGGTCTTTATTA 780
721 GGGGAACCAATTTCTGCAAAACCCAGATGGGAGCAAAAGCCCTGGTCTTTATTA 780
781 GTTACCAATGACAAGTGAATGGAAATCTGTAGTCTGACGTGCTGACCCAGAC 840
781 GTTACCAATGACAAGTGAATGGAAATCTGTAGTCTGACGTGCTGACCCAGAC 840
841 GTTGGCTCCCAAGGAAACCCCACTGAGCCATCAACCAAGCTTCCGGGTTTGAACCTC 900
841 GTTGGCTCCCAAGGAAACCCCACTGAGCCATCAACCAAGCTTCCGGGTTTGAACCTC 900
901 TGTGAAAGACTGAGATAGCAGAGAGAAAGATCAAGAGATCTATGGAGCTTTAAGAC 960
901 TGTGAAAGACTGAGATAGCAGAGAGAAAGATCAAGAGATCTATGGAGCTTTAAGAC 960
961 ACGGGGGCAAGCCCATGGAGGAGGCTCCCTCCAGTCTGCTGACCATCTCC 1020
961 ACGGGGGCAAGCCCATGGAGGAGGCTCCCTCCAGTCTGCTGACCATCTCC 1020

961 ACGGGGGCAAGCCCATGGAGGAGGCTCCCTCCAGTCTGCTGACCATCTCC 1020
1021 ATGCCCAAGGGCCACTTCTGTGTGTGGGCGCTGATTCACCCCTGCTGGGTGCTACAGCT 1080
1021 ATGCCCAAGGGCCACTTCTGTGTGTGGGCGCTGATTCACCCCTGCTGGGTGCTACAGCT 1080
1081 GCCCAGTGCACCCGATTAATAAACAGACATCTAAGGTGTGTGTAGGGGACAGGACCTG 1140
1081 GCCCAGTGCACCCGATTAATAAACAGACATCTAAGGTGTGTGTAGGGGACAGGACCTG 1140
1141 AAGAAAGAGATTTTATGATAGCAGAGCTTTAGGGTGGAGAAAGATTAATCAAGTACAGCCAC 1200
1141 AAGAAAGAGATTTTATGATAGCAGAGCTTTAGGGTGGAGAAAGATTAATCAAGTACAGCCAC 1200
1201 TACATGAAAGAGATGAGATTCGCCCAATGATATTGCTATTGCTCAAGTTAAAGCCACTG 1260
1201 TACATGAAAGAGATGAGATTCGCCCAATGATATTGCTATTGCTCAAGTTAAAGCCACTG 1260
1261 GATGCTACAGTGTCTAGAAATCCAAATACGTAAGAGCTGTGTGCTGATGAGTCC 1320
1261 GATGCTACAGTGTCTAGAAATCCAAATACGTAAGAGCTGTGTGCTGATGAGTCC 1320
1321 TTTCCCTGTGGAGTGAATGCCACATCTCTGCGGTGGTGTACAGAAACAGGAAAGG 1380
1321 TTTCCCTGTGGAGTGAATGCCACATCTCTGCGGTGGTGTACAGAAACAGGAAAGG 1380
1381 TCCCGGCACTCTCTGATGCGCAAGTCAAGCTGATGCGCAAACTTTGTGCAACTCCCGC 1440
1381 TCCCGGCACTCTCTGATGCGCAAGTCAAGCTGATGCGCAAACTTTGTGCAACTCCCGC 1440
1441 CAACCTATGACACATGATGATGACATGATGATGATGATGATGATGATGATGATGATGAT 1500
1441 CAACCTATGACACATGATGATGACATGATGATGATGATGATGATGATGATGATGATGAT 1500
1501 GGCGAAGACCTTCCAGGCTGACTGTGAGGCGCCCTGACCTGTGAGAAAGGACGCGAC 1560
1501 GGCGAAGACCTTCCAGGCTGACTGTGAGGCGCCCTGACCTGTGAGAAAGGACGCGAC 1560
1561 TACTACGTCTATGAGGATGATGAGCTGGGCTGAGTGTGGAAAGACCGAGGGTCTAC 1620
1561 TACTACGTCTATGAGGATGATGAGCTGGGCTGAGTGTGGAAAGACCGAGGGTCTAC 1620
1621 ACCCAAGTACCAAAATCTGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
1621 ACCCAAGTACCAAAATCTGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
1681 TAA 1683
1681 TAA 1683

RESULT 4
US-09-864-761-11164/c
Sequence 11164, Application US/09864761
Patent No. US2002048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmiga-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04

```

; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonex Sequence Listing Engine vers. 1.1
; SEQ ID NO 11164
; LENGTH: 451
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006097.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.89
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.9
US-09-864-761-11164

```

```

Query Match 7.1%; Score 120.2; DB 10; Length 451;
Best Local Similarity 97.6%; Pred. No. 4.9e-27;
Matches 122; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 324 TCAGAAAGTGCAGAAATACGTGCAGAGACACCATGTGCGCGGCCCAATGTCATTAC 383
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 451 TCCTACAGTGCAGAAATACGTGCAGAGACACCATGTGCGCGGCCCAATGTCATTAC 392
QY 384 CCAGAGTCCTCCCTACCTACCGCTGTCTGTAAACACCTTAAACAGTCCAGCTGTC 443
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 391 CCAGAGTCCTCCCTACCTACCGCTGTCTGTAAACACCTTAAACAGTCCAGCTGTC 332
QY 444 CCAG 448
    |||||
DB 331 CCAG 327

```

```

RESULT 5
; Sequence 552: Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617

```

```

; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 552
; LENGTH: 2036
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-552

```

```

Query Match 7.0%; Score 117.6; DB 10; Length 2036;
Best Local Similarity 49.8%; Pred. No. 8.2e-26;
Matches 607; Conservative 0; Mismatches 559; Indels 54; Gaps 10;

```

```

QY 469 AACCCCTCCGAGATGGGCTACCTGCTCCGGCATTAAGCGGATCCAAATTCACCGT 528
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 745 ACCCTTCCTCGAAGCGGGGACACCTGCCATGTCGTGGCACCGGACACCCGTGT 804
QY 529 GCCCTGCCGACAGTTCAAGGAAATTCGTGAATAGTTTGATCA---CTGCTAT 585
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 805 GCCCTGCCGACAGCTTCGTGTGAGCGGCTGTGCAACTCGAGCTGATGAGCGCTGC 864
QY 586 GTTGGCGATGCTACTCTTACGAGAGGAAATGAAATGAGACATCMAACGACATGCTGC 645
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 865 TTGGGGAACGGCACTGTGATACGTTGCGTGCGGCGGACGACCTCAGCTCGGCGCTCAGCTGC 924
QY 646 CTTTACGTGACATCCACCTCTCTTTCAGAGAGAAATTACAACATCTTTATGAGATGCT 705
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 925 CTGGCTTGAATCTCGATCTGTCTTACAGAGAGCTGCACTGAGATCTGTTGGCGCGCG 984
QY 706 GAACCCATGGGATTTGGGGAACACAAATTTCTGCAGAAACCCAGATCCGAGCAAAAGCCC 765
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 985 GCCCTGTGGGCTGGGCGGCCCATGCTTACTGCGGAATCCGGCAATGACAGAGGCC 1044
QY 766 TGGTCTTTATTAAGTTTACATGACAAAGGTGAATGGGAATACGTGTCAAGCC 825
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1045 TGGTCTTACGT---GTTGAGAGACAGCGGCTCTCTGAGATACGCGCTGAGAGCC 1101
QY 826 TGTCTAGCCAGGACGTTGCTTACCCAGAGAAAGCCCACTGAGCATCAACCAAGCTT 885
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1102 TCGAATCCCT-----CACCAAGTCCAACTGTACCGGATCTCTGCGACCTG 1152
QY 886 CCGGGTTTGAATCTCTGTGGAAGAGCTGATAGACAGAGAAATCAAGAGATCTAT 945
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1153 CCTGAGCCAGCTCCCGGGGGGCCAGAGCTGCGGCAAGAGACAAAGAGAGAGCTTC 1212
QY 946 GAGGCTTTAAGAGACAGCGGCGGCAAGACCCATGCGAGCGCTCTCAGTCTCGCTG 1005
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1213 CTGGCGCCAGTATCATGCGGCTCTCTCTGCGCGGCTGCGACCCCTG----- 1265
QY 1006 CCTGACCATCTTCATGCCCGGAGCACTTCTGTGTGGGGCGGTATCCACCGCTGC 1065
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1266 --GCTGGCCCATCTCATGCGGAGACGCTTGGCGCGGAGCTGTGCACACCTGC 1223
QY 1066 TGGGTGCTACTGTGCTGCCCTGCAACCGACATTAACCAACAGACATCTTA-----AGTGG 1119
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db	1324	TGGGTGTGTGTGGCGGGCCCACTGCTTCTCCCAACACCCCCCAAGGACACGCTTCCTGGT	1383
QY	1120	GTGCTTAGGGGACCAAGACCTGAAAGAAAGAATTTATYAGCAGAGCTTTAGGTGGAG	1179
Db	1384	GTTCCTGGGGCCAGACACTTCTTTCAACGCGACGACGACGTGACGAGACTTTTGGCAATCAG	1443
QY	1180	AAGATATTCAAGTACAGCCACTCAATATGAAACAGATGAGATTTCCCAACATGATATTGCA	1239
Db	1444	AAGTATACCTCCGTACACCTGTACTGCGTGTTCACCCACAGCCAGC---GACCTCGTC	1500
QY	1240	TTGCTCAAGTTAAAGCCAGTGCATGGTCACTGTGCTCTAGATTCCAATTAATCGTGAAGACT	1299
Db	1501	CTGATCTCGGCTGAAGAAAGAAAGGGAGCGGTGTGCCACAGCTTGCAATTTCGACGCC	1560
QY	1300	GTGTGCTTTCCCTGAT-----GGGTCTTTCCCTCTGGGAATGAGTCCACATCTCTGGC	1353
Db	1561	ATCTGCTCTCCGACGCCGAGACACTTCCCGAGGACCAAGTCCAGATATGGCGGC	1620
QY	1354	TGGGG-----TGTTACGAGAAACGAGGAAAGGTCCGCGACGTCTCGATGTGCGAAAGTC	1407
Db	1621	TGGGGCACTTGATGATGAGACAGTGAAGGCTACTCCACTCCCTCGGGAAGGCGCTTGCTC	1680
QY	1408	AAGCTGATTTGCCACACACTTTGTGCACTCCCGCACACTCTATGACCAACATGATGATGAC	1467
Db	1681	CCCTGTGTCGGCGACACACAAATGTGACGAGCCTTGAGTCTTCAGCGCGCGACATCAACGCC	1740
QY	1468	AGATATGCTGTGACGAGAAATCTTCACAAACCTGGGCAAGACACTCCAGGGGTGACTCT	1527
Db	1741	AACATCTCTGTGGCGGCTACTT---CGACTGCAATGTCCAGGCTTCCAGGGGAGCTCA	1797
QY	1528	GGAGGCCCTTCAGCTGTGAGAGAGAGCGCACTTACTACGTCTATGAGGATAGTGAACCTGG	1587
Db	1798	GGGGGGGCCCTTGCTCTGCGAAGAAACAGGGGTGGCTTACTCTTCAGGCAATCAATCAACTGG	1857
QY	1588	GGCCTGGAGCTGGGG-----AAGAGCCAGGGGGTCTACACCAACGAATTCACATTTCTTG	1641
Db	1858	GGTGAAGGGCTGGGGCGGCTCCACAAACCGGGGGTCTACACCGCGTGGCCCACTATTGG	1917
QY	1642	AATTTGGATCAAAAGCCACAT	1661
Db	1918	GACTGGATCAACGACCGGAT	1937

```

US-09-880-107-1612
? Sequence 1612, Application US/09880107
? Patent No. US20020142981A1
? GENERAL INFORMATION:
? APPLICANT: Horne, Darci T.
? APPLICANT: Vockley, Joseph G.
? APPLICANT: Scherf, Uwe
? APPLICANT: Gene Logic, Inc.
? TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
? FILE REFERENCE: 44921-5028-WO
? CURRENT APPLICATION NUMBER: US/09/880,107
? CURRENT FILING DATE: 2001-06-14
? PRIOR APPLICATION NUMBER: US 60/211,379
? PRIOR FILING DATE: 2000-06-14
? PRIOR APPLICATION NUMBER: US 60/237,054
? PRIOR FILING DATE: 2000-10-02
? NUMBER OF SEQ ID NOS: 3950
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 1612
? LENGTH: 2036
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? OTHER INFORMATION: Genbank Accession No. US20020142981A1 D141012
US-09-880-107-1612

```

Query Match	7.08;	Score 117.6;	DB 10;	Length 2036;
Best Local Similarity	49.88;	Pred. No. 8.2e-26;		
Matches 607; Conservative	0;	Mismatches 559;	Indels 54;	Gaps 10;

QY	469	AACCTCTGCCAAGTAATGGGGCTACGTCGTCGCCGATAAAGGGAGATCCACAGTCCACTGT	528
Db	745	AAGCCTTTGCTCTGAACGCGGGGACACTGCCACTGATGTGGCCACCGGGACCAACGGTGTGT	804
QY	529	GGCTCTCCGACACGATTCAAGGGGAAATTTCTGTCAATTAAGTGTCTGATGA---CTGCTAT	585
Db	805	GGCTCCGCCACACAGGCTTCGCTGGAGGGCTCTGGCAACATCGAGCTGAATGACGCTGTCTC	864
QY	586	GTTCGCGATGAGCTCTACTTACACGAGGGAAATTAAGATGGAACAGTCAACACCATGCTGTGC	645
Db	865	TTGGGGGAACGCACTGGGTATCCCTGTCCGTGGCCGACGACCTCACCTCGGGCCCTCAGCTGC	924
QY	646	CTTTACTGTGAACCTCCCACTCCCTCTTTCGAGAGAGAAATTAACAACATGTTTATGAGAGATGCT	705
Db	925	CTGGCCTCGGAAGCTCCGATCTGCTCTTACAGAGACTGCACGTGACCTCCGTGGGGCGCGC	984
QY	706	GAACCCCATGGGATTTGGGGAAACAATTTCTGCGAAAACCCAGATGCGAGAAAGCC	765
Db	985	GGCTCGCTGGGGCCCTGGGGCCCCCTATGCTACTGTCCGGAAATCCGAGCAATGACAGAGGCC	1044
QY	766	TGGTCTTTTATTAAGTTATCAATGACAAAGGTGAAGAAATGGGAATATGATGTCTCAACC	825
Db	1045	TGGTCTGATAGT---GGTGAAGAACAGCGGGCTCTCTGGGAGTACTGCCGCTGGAGGCC	1101
QY	826	TGCTACGCCCAAGAGACGTGTGCTATCCACAGGAAAACCCCACTAGCATCAACCAAGCTT	885
Db	1102	TGGGAATCCCT-----CACAGAGTCCACTGTACCGGATCTCTTGGGCAACCTG	1152
QY	886	CCGGGGTTTGACTCCTGTGGAAAGACTGAGATGACAGAGAGAAATCAAGAAATCTAT	945
Db	1153	CTGAGACCAAGCTTCCCGGGGGGCCAGGGCTCTGCGGACAGGAGGACAAAGAGAGGACGTTT	1212
QY	946	GGAGGCGTTTAAGACACAGGAGGGGCAAGCCATGAGCAAGCGTCCCTCAGTCCGCGTG	1005
Db	1213	CTCGGGCAACGTATCATGTGGGGGCTCTCTCTGCTGCCCGGCTCGCACCCCTG-----	1265
QY	1006	CCCTGTACCATCTGCATATGCCCAAGGGCCACTTCTGTGTGGGGGCGTGTATCCACCCCTGC	1065
Db	1266	--GCTGGCGGCATCTACATGTGGGGACACCTTCTGGCGCGGAGCCTGGTCTCACACCTGC	1323
QY	1066	TGGGTGCTACTGCTGTGCCACTGCACCCGACATAAAAACGAGACATCTAA-----AGGTG	1119
Db	1324	TGGTGTGTGTGGCGCGGCCCTCACTGCTTCTCCACACCCCCCAAGGGACAGCTGTCCCTG	1383
QY	1120	GTCCTAAGGGGAAACAGAGACTCGAAGAAAGAAATTTCTATGAGCAGAGCTTTAGGGTGAAG	1179
Db	1384	GTCGTGGGCGACACTTCTTCAACCGCAGACGAGACGTGAGCGAGACCTTGGGCAATCGAG	1443
QY	1180	AAGATATTCAAGTACAGCCACTCAATGAAGAAGATAGATTTCCCAACAATATATTTGCA	1239
Db	1444	AAGTATACATCCGCTACACCTGTACTCGGTTCACACCCCAACGACGACA---GACCTGCTC	1500
QY	1240	TTGCTCAAGTTAAACCCAGATGAGTGTGCTACTGTGCTAGAAATCCAAATATCGTGAAGACT	1299
Db	1501	CTGATTCGGCTCAAGAAAGAAAGGGAGCCCTGTGCGCACACCTTCGCAAGTTCTGCGAGCCC	1560
QY	1300	GTCGTGTATGCCGAT-----GGGTCTTCTTCTCTGTGGAGTGAAGGCCCAACTCTCTGGC	1353
Db	1561	ATCTGCTCTCCCGAAGCCGGCAGCACCCTTCCCCGAGACACAAGATGCCAGTATGCGGGC	1620
QY	1354	TGGGG-----TGTTACAGAAACAGGAAAGGGTCCGCCAGCTCCTGGATGCAAAAGTC	1407
Db	1621	TGGGGCCACTTTGGATGAGAAAGGTGACGGGCTACTCACACTCCCTGGCGGAGGCCCTGTGTC	1680
QY	1408	AAGCTATTGGCAACACTTTTGTGCACACTCCCCCAACTCTGTACACCAATGATTTGATGAC	1467
Db	1681	CCCTCGTGTGGCGAACCAAGGTGAGAGCCCTGAGGTCTTACGGCGGCCGACATCAGACCCC	1740
QY	1468	AGTATATCTGTGACAGAAATCTTTCAGAAACCTGGGCAAGACACTGTGCCAGGGTGTACTCT	1527
Db	1741	AAGATCTCTGTGGCGGCTACTT---CGATCCAAAGTCCGAGGCTCTCCACAGGGGAGCTCA	1797


```

Db 1367 CTCCTTCTATTGAGAGCGGCTGAGAGGCGCTATGCTAGCTATGACCATCCACCGCCT 1426
QY 1430 GCACCTCCCGCCCACTCTATGACCATGATTTGATGACATGATGATGTCGAGAAATC 1489
Db 1427 GCACATACACAACTTACTTAAACAGAGTCACCGCAACACATGCTGTGTGTGAGACA 1486
QY 1490 TTCAGAACTG-----GCAGACACTGTCAGGCTGACTCTGTGAGGCC 1534
Db 1487 CTCGAGCGCGGCGCCCGCCAGCAACTTTCAGACGACCTGCGAGGCGCATTTGAGAGCC 1546
QY 1535 CCCGACCTGTGAGAGGAGCGACCTACTGCTATGAGATGATGAGCTGGGCGCTGG 1594
Db 1547 CCGTGTGTGTGAGACGATGAGCGCCATGACTTGTGTGATCATCATCAGCTGGGCGCTGG 1606
QY 1595 AGGTGGAGAGG-----CCAGGGGCTCTACACCAAGTTACCAATTCCTGAATTGGA 1648
Db 1607 GCTGTGACAGAGAGATGTCGCGGTGTGTACACCAAGTTACCAACTGAGACTGGA 1666
QY 1649 TCAAGCCACCAT 1661
Db 1667 TTCGTGACACAT 1679

```

RESULT 9

```

US-09-969-271-5
; Sequence 5, Application US/09969271
; Patent No. US20020098179A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (SB));
; APPLICANT: Pfizer Limited (GB and EP (GB) only)
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PCS10951AYPE
; CURRENT FILING DATE: US/09/969, 271
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 0025473.0
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2519
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-271-5

```

```

Query Match 4.5%; Score 75.4; DB 10; Length 2519;
Best Local Similarity 50.8%; Pred. No. 1.2e-12;
Matches 342; Conservative 0; Mismatches 286; Indels 45; Gaps 5;

```

```

QY 1028 AGGGCCACTTCTGTGTGGGGCCCTGATCCACCCTGCTGGGTCTGCTAGCTGCCACT 1087
Db 1089 AGCGGTTCCTGTGGGGGGCATCATCACTCTGCTGTGATTCCTGTGCGGCCACT 1148
QY 1088 GCACCGA-----CATAAACAGACATCTAAAGTGTGCTAGGGAGACAGACCTGA 1141
Db 1149 GCTTCAGAGAGAGTTCCGCCACCACTGAGGCTGATTTGGCAGAACTACCGGG 1208
QY 1142 AGAAGAAGATTTTCATGACGACAGACTTAGGGTGAAGAATATTTCAATATAGCCACT 1201
Db 1209 TGGTCCCTGGCGAGAGAGCAAAATTTGAAGTGAATAATTCATTTGCCATTAAGAAAT 1268
QY 1202 ACAATGAAGAAGATGATTCACCAATGATATTTGCACTTTCACCAATTAAAGCCAGTGG 1261
Db 1269 TCGAT-----GATGACACTTACGACATGACATTTGGCTGCTGACACTGAATTCGATTT 1322
QY 1262 ATGTGCTACTGTGCTCTAGATCAATACGTGAAGACTGTGTCTGCTGCTGATAGGGTCT 1321
Db 1323 CGTCCCGCTGTCCAGAGAGAGCGCTGTCCGACATGTGTGCTTCCCGCGCGGAGCC 1382
QY 1322 TTCCT-----CTGGGAGTGAAGTGCACATCTGCGCTGGGAGTTACAGAAA 1369
Db 1383 TGGAGTGTCCGAGACTGAGCGAGAGTGTGAGCTCTCGGCTACGCGCAAGCATAGAGCCCTTGT 1442
QY 1370 CAGGAAAGGCTCCCGCAGCTCTCTGATGCCAAAGTCAAGCTGATTTGCCAACCTTTGT 1429

```

```

Db 1443 CTCCTTCTATTGAGAGCGGCTGAGAGGCGCTATGCTAGCTATGACCATCCACCGCCT 1502
QY 1430 GCACCTCCCGCCCACTCTATGACCATGATTTGATGACATGATGTCGAGAAATC 1489
Db 1503 GCACATACACAACTTACTTAAACAGAGTCACCGCAACACATGCTGTGTGTGAGACA 1562
QY 1490 TTCAGAACTG-----GCAGACACTGTCAGGCTGACTCTGTGAGGCC 1534
Db 1563 CTCGAGCGCGGCGCCCGCCAGCAACTTTCAGACGACCTGCGAGGCGGATTCGAGAGCC 1622
QY 1535 CCCGACCTGTGAGAGGAGCGACCTACTGCTATGAGATGATGAGCTGGGCGCTGG 1594
Db 1623 CCGTGTGTGTGAGACGATGAGCGCCATGACTTGTGTGATCATCATCAGCTGGGCGCTGG 1682
QY 1595 AGGTGGAGAGG-----CCAGGGGCTCTACACCAAGTTACCAATTCCTGAATTGGA 1648
Db 1683 GCTGTGACAGAGAGATGTCGCGGTGTGTACACCAAGTTACCAACTGAGACTGGA 1742
QY 1649 TCAAGCCACCAT 1661
Db 1743 TTCGTGACACAT 1755

```

RESULT 10

```

US-09-974-298-144
; Sequence 144, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Huel-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974, 298
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/258, 331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 144
; LENGTH: 2641
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020156263A1 1001470CBI
; LOCATION: 2635
; OTHER INFORMATION: a, t, c, g, or other
US-09-974-298-144

```

```

Query Match 4.5%; Score 75.4; DB 9; Length 2641;
Best Local Similarity 50.8%; Pred. No. 1.2e-12;
Matches 342; Conservative 0; Mismatches 286; Indels 45; Gaps 5;

```

```

QY 1028 AGGGCCACTTCTGTGTGGGGCGCTGATCCACCCTGCTGGGTCTGCTAGCTGCCACT 1087
Db 1209 AGCGGTTCCTGTGGGGGGCATCATCACTCTGCTGTGATTCCTGTGCGGCCACT 1268
QY 1088 GCACCGA-----CATAAACAGACATCTAAAGTGTGCTAGGGAGACAGACCTGA 1141
Db 1269 GCTTCAGAGAGAGTTCCGCCACCACTGAGGCTGATTTGGGAGCAACATACCGGG 1328
QY 1142 AGAAGAAGATTTTCATGACGACAGACTTAGGGTGAAGAATATTTCAATATAGCCACT 1201
Db 1329 TGTCCCTGGCGAGAGAGAGCAAAATTTGAAGTGAATAATTCATTTGCCATTAAGAAAT 1388
QY 1202 ACAATGAAGAAGTGAATTCACCAATGATATTTGCACTTTCACCAATTAAAGCCAGTGG 1261
Db 1389 TCGAT-----GATGACACTTACGACATGACATTTGGCTGCTGACACTGAATTCGATTT 1442
QY 1262 ATGTGCTACTGTGCTCTAGATCAATACGTGAAGACTGTGTGCTGCTGCTGATAGGGTCT 1321
Db 1443 CGTCCCGCTGTCCAGAGAGAGCGTGTGCGCACTGTGTGCTTCCCGCGCGGAGCC 1502

```


PRIOR FILING DATE: 1998-04-09	PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-03	PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15	PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15	PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15	PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15	PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15	PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21	PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21	PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22	PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22	PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22	PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22	PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23	PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27	PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-26	PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083429
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083544
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083555
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083555
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083555
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083744
PRIOR FILING DATE: 1998-04-30	PRIOR APPLICATION NUMBER: 60/083366
PRIOR FILING DATE: 1998-05-05	PRIOR APPLICATION NUMBER: 60/084411
PRIOR FILING DATE: 1998-05-06	PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06	PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07	PRIOR APPLICATION NUMBER: 60/084633
PRIOR FILING DATE: 1998-05-07	PRIOR APPLICATION NUMBER: 60/084644
PRIOR FILING DATE: 1998-05-07	PRIOR APPLICATION NUMBER: 60/084599
PRIOR FILING DATE: 1998-05-07	PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07	PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07	PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07	PRIOR APPLICATION NUMBER: 60/085333
PRIOR FILING DATE: 1998-05-13	PRIOR APPLICATION NUMBER: 60/085333

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 12:17:47 ; Search time 4285 Seconds

(without alignments)
11430.577 Million cell updates/sec

Title: US-09-912-559-2

Perfect score: 1683

Sequence: 1 atgttcgacagatgtctga.....aaagtgaagtgtctctaa 1683

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenBank:*

1: gb_da:*

2: gb_hlg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_da:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vi:*

30: em_hlg_hum:*

31: em_hlg_inh:*

32: em_hlg_other:*

33: em_hlg_mus:*

34: em_hlg_pln:*

35: em_hlg_rod:*

36: em_hlg_mam:*

37: em_hlg_vrt:*

38: em_sy:*

39: em_hlgo_hum:*

40: em_hlgo_mus:*

41: em_hlgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1683	100.0	1683	6	AX383955
2	1679.8	99.8	1683	6	AX383954
3	1679.8	99.8	2408	6	S83182
4	1679.8	99.8	3008	6	AX409021
5	1679.8	99.8	3008	9	HUMHGFAL
6	1678.2	99.7	2251	9	BC031412
7	1047	62.2	2119	10	BC031775
8	261.2	15.5	188349	9	AL390197
9	258	15.3	157750	2	AC006097
10	241	14.3	397	6	AX305372
11	167.4	9.9	200960	2	AC115771
12	129.8	7.7	2035	10	AF099017
13	129.8	7.7	2063	10	AF224724
14	128.2	7.6	2067	10	BC019376
15	119	7.1	68586	2	AC106236
16	114.4	6.8	2033	6	E07591
17	114.4	6.8	2033	6	I15469
18	114.4	6.8	2033	6	I69315
19	114.4	6.8	2036	6	AX33070
20	114.4	6.8	2036	6	AX408965
21	114.4	6.8	2036	9	HUMHGFAP
22	99.8	5.9	2228	4	BTUPA
23	99.8	5.9	2355	4	BOVUPKA
24	99	5.9	2375	4	SSUPAR
25	94.6	5.6	1454	10	RNUPAM
26	94.6	5.6	2366	10	RNUPPLAC
27	93.2	5.5	2299	10	MMURKR
28	93	5.5	970	6	E07615
29	93	5.5	970	6	I15462
30	93	5.5	970	6	I69308
31	87.6	5.2	2245	4	VMPPA1A
32	87.6	5.2	2252	6	A07242
33	87.6	5.2	2252	6	A07250
34	87.6	5.2	2266	4	VMPPA1S
35	87.4	5.2	2512	6	A19618
36	87.2	5.2	2445	10	RATPATISS
37	84.8	5.0	2130	4	VMPAB
38	84.8	5.0	2137	6	A07246
39	81.2	4.8	1620	6	I08184
40	81.2	4.8	1725	4	VMPTPA
41	81.2	4.8	2236	4	VMPPA2A
42	81.2	4.8	2257	6	A07244
43	77.6	4.6	2460	4	AC364605
44	77.2	4.6	919	6	A93672
45	77.2	4.6	2393	5	CHKUPAR

ALIGNMENTS

RESULT 1

AX383955

LOCUS AX383955 1683 bp DNA linear PAT 19-MAR-2002

DEFINITION Sequence 2 from Patent EP1182258.

ACCESSION AX383955

VERSION AX383955.1 GI:19577526

KEYWORDS

SOURCE

ORGANISM

human.

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 Roemisch,J., Stoehr,H.A., Feussner,A., Lang,W., Weimer,T.,

Becker,M., Nerlich,C. and Muth-Naumann,G.

Mutants of the factor vii activating protease and methods for their

Pred. No. is the number of results predicted by chance to have a

detection
Patent: EP 1182258-A 2 27-FEB-2002;
Aventis Behring GmbH (DE)

FEATURES
Location/Qualifiers
Source 1..1683

/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 441 a 437 c 435 g 370 t
ORIGIN

Query Match 100.0%; Score 1683; DB 6; Length 1683;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1683; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 AVGTTTCCAGAGATGCTGATCTCCATGTTCTGCTGTTAATGGCTCTGGTGGAGAACA 60
DB 1 ATGTTTCCAGAGATGCTGATCTCCATGTTCTGCTGTTAATGGCTCTGGTGGAGAACA 60
OY 61 GCCGTGGGTTCTCCCTGATGCTTTATGGAAGCCTGGACCCAGACTGGACCCCTGGAC 120
DB 61 GCCGTGGGTTCTCCCTGATGCTTTATGGAAGCCTGGACCCAGACTGGACCCCTGGAC 120
OY 121 CAGTATGATTACAGCTACAGAGATTATATCAGAGAGAACAGCAGTACAGCACTTACC 180
DB 121 CAGTATGATTACAGCTACAGAGATTATATCAGAGAGAACAGCAGTACAGCACTTACC 180
OY 181 CAGTATGAGATCTGACCTGAGTACAGCAGGAGACCACTGATCCATGCCAGCCAC 240
DB 181 CAGTATGAGATCTGACCTGAGTACAGCAGGAGACCACTGATCCATGCCAGCCAC 240
OY 241 CCCTGTGAACAGCGTGGGAGCTGCCGTGCATGGAGACACCTTGACATGACGCTGCTG 300
DB 241 CCCTGTGAACAGCGTGGGAGCTGCCGTGCATGGAGACACCTTGACATGACGCTGCTG 300
OY 301 GCTCCTTTCTGGGGAATTAAGTGTGAGAAAGTGAATACGTGAGAGAGACCACTGAT 360
DB 301 GCTCCTTTCTGGGGAATTAAGTGTGAGAAAGTGAATACGTGAGAGAGACCACTGAT 360
OY 361 GGGCGGGGCAATGCTCATTAACCAAGATCCCAAGTTCACTGCTGTGTGTAACAC 420
DB 361 GGGCGGGGCAATGCTCATTAACCAAGATCCCAAGTTCACTGCTGTGTGTAACAC 420
OY 421 CCTTACACAGAGTCCAGCTGCTCCAGAGTGTCTGTATGAGAGCCCAACCCCTGGCAG 480
DB 421 CCTTACACAGAGTCCAGCTGCTCCAGAGTGTCTGTATGAGAGCCCAACCCCTGGCAG 480
OY 481 AATGGGCTTACTGCTCCCGCATTAAGGAGATCCCAAGTTCACTGCTGTGTGTAAC 540
DB 481 AATGGGCTTACTGCTCCCGCATTAAGGAGATCCCAAGTTCACTGCTGTGTGTAAC 540
OY 541 CAGTTCAAGGGGAAATTTCTGTAAGTGTGATGATGCTGATGTTGGAGATGGCTTAC 600
DB 541 CAGTTCAAGGGGAAATTTCTGTAAGTGTGATGATGCTGATGTTGGAGATGGCTTAC 600
OY 601 TCTTACCGAGGGAATTAATGAGACAGTCAACAGCATGCTGCTGCTTACTGGAAC 660
DB 601 TCTTACCGAGGGAATTAATGAGACAGTCAACAGCATGCTGCTGCTTACTGGAAC 660
OY 661 CACCTCCTCTGAGAGAAATTAACACATGTTATGAGAGATGTAAGCAACCATGGGATT 720
DB 661 CACCTCCTCTGAGAGAAATTAACACATGTTATGAGAGATGTAAGCAACCATGGGATT 720
OY 721 GGGGAACACAAATTTCTGCAAGAACCCAGATGCGGAGAAAGCCCTGCTGTTAATAA 780
DB 721 GGGGAACACAAATTTCTGCAAGAACCCAGATGCGGAGAAAGCCCTGCTGTTAATAA 780
OY 781 GTTACCAATGACAGGTAATGGAATGGAATGATGATGCTGCTGCTGCTGCTGCTGCTG 840
DB 781 GTTACCAATGACAGGTAATGGAATGGAATGATGATGCTGCTGCTGCTGCTGCTGCTG 840
OY 841 GTTGCTTACCCAGAGAGAACCCCACTGAGCCATCAACCAAGCTTCCGGGGTTGACTCC 900
DB 841 GTTGCTTACCCAGAGAGAACCCCACTGAGCCATCAACCAAGCTTCCGGGGTTGACTCC 900
```

```
OY 901 TGTGGAAGAGCTGATAGCAGAGAGGAAGATCAAGAGATCTATGAGGCTTTAAGAC 960
DB 901 TGTGGAAGAGCTGATAGCAGAGAGGAAGATCAAGAGATCTATGAGGCTTTAAGAC 960
OY 961 ACCGGGGCAAGCAGCCCATGAGGAGGCTCCCTCCAGTCTGCTGCTGCTGAGCATCTCC 1020
DB 961 ACCGGGGCAAGCAGCCCATGAGGAGGCTCCCTCCAGTCTGCTGCTGAGCATCTCC 1020
OY 1021 ATGCCCAAGGCGCATCTCTGCTGGTGGGCGCTGATCAACCCCTGCTGCTGCTGCT 1080
DB 1021 ATGCCCAAGGCGCATCTCTGCTGGTGGGCGCTGATCAACCCCTGCTGCTGCTGCT 1080
OY 1081 GCCCATGCTACCGCATTAATAAACCAAGATCTAAAGTGTGCTAGAGGAGCAGAGACCTG 1140
DB 1081 GCCCATGCTACCGCATTAATAAACCAAGATCTAAAGTGTGCTAGAGGAGCAGAGACCTG 1140
OY 1141 AAGAAAGAGAAATTTCAATGAGCAGAGCTTTAGGGTGCAAGATATTCAAGTACAGCCAC 1200
DB 1141 AAGAAAGAGAAATTTCAATGAGCAGAGCTTTAGGGTGCAAGATATTCAAGTACAGCCAC 1200
OY 1201 TACAATGAAGAGATGATGATTTCCCAATGATATGATGCTGCTCAAGTTAAAGCAGTG 1260
DB 1201 TACAATGAAGAGATGATGATTTCCCAATGATATGATGCTGCTCAAGTTAAAGCAGTG 1260
OY 1261 GATGCTCATGCTGCTCTAGATCAATTAACGTGAAGACTGTGTGCTTGCCTGATGGCTC 1320
DB 1261 GATGCTCATGCTGCTCTAGATCAATTAACGTGAAGACTGTGTGCTTGCCTGATGGCTC 1320
OY 1321 TTTCCCTCTGGAGTGAAGTCCACATCTCTGCTGGGCTGTACAGAAACAGAGAAAGG 1380
DB 1321 TTTCCCTCTGGAGTGAAGTCCACATCTCTGCTGGGCTGTACAGAAACAGAGAAAGG 1380
OY 1381 TCCGGCCAGCTCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1440
DB 1381 TCCGGCCAGCTCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1440
OY 1441 CAACTCTATGACCATGATTTGATGACAGTATGATCTGTGAGAGAAATCTTCAGAAACCT 1500
DB 1441 CAACTCTATGACCATGATTTGATGACAGTATGATCTGTGAGAGAAATCTTCAGAAACCT 1500
OY 1501 GGGCAAGACACTGCGGAGGAGTCTGAGAGGCGCCCGGACCTGTGAGAGAGAGAGAG 1560
DB 1501 GGGCAAGACACTGCGGAGGAGTCTGAGAGGCGCCCGGACCTGTGAGAGAGAGAGAG 1560
OY 1561 TACTAGCTATGAGGATGAGCTGGGCTGAGAGTGTGAGAGAGAGAGAGAGAGAG 1620
DB 1561 TACTAGCTATGAGGATGAGCTGGGCTGAGAGTGTGAGAGAGAGAGAGAGAGAG 1620
OY 1621 ACCCAAGTTACCAATCTGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 1680
DB 1621 ACCCAAGTTACCAATCTGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 1680
OY 1681 TAA 1683
DB 1681 TAA 1683
```

RESULT 2
AX383954
LOCUS 1683 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 1 from Patent EP1182258.
ACCESSION AX383954
VERSION AX383954.1 GI:19577525
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
Roemisch, J., Stoeck, H.A., Feussner, A., Lang, W., Weimer, T.,
Becker, M., Nerlich, C., and Muth-Naumann, G.
Mutants of the factor vii activating protease and methods for their

TITLE Purification and characterization of a novel hyaluronan-binding protein (PHBP) from human plasma: it has three EGF, a kringle and a serine protease domain, similar to hepatocyte growth factor activator

JOURNAL U. Biochem. 119 (6), 1157-1165 (1996)

MEDLINE 96425001

PUBMED 8827452

REMARK Genbank staff at the National Library of Medicine created this entry [NCBI 91bdsq 179630] from the original journal article. This sequence comes from Fig. 3B.

FEATURES

Source Location/Qualifiers

1..2408

/organism="Homo sapiens"

/db_xref="taxon:9606"

1..2408

/gene="hyaluronan-binding protein, PHBP"

97..1779

/gene="hyaluronan-binding protein, PHBP"

/note="hepatocyte growth factor activator homolog; Method: conceptual translation with partial peptide sequencing; This sequence comes from Fig. 3B; PHBP"

/codon_start=1

/product="hyaluronan-binding protein"

/protein_id="AAB46909.1"

/db_xref="GI:1836159"

/translation="MFARMSDLHVLMLALVGTACGFSLMSLESLDPDWTPEQDYD SYEDNOEENTSTLHAENDPMVYTEDQADFCQNPENHSGDLVSGVFTSCILAP FSGNCKOVMTCKDNPCGSGOCLITQSPYRCCKHPYRGKMYVQVVCNPNQCO NGATCSRKRSPKFTACPDQFKECEIGSDQCVGPGYGYRKNRMRYVNOIACLYM NSHLLLOENYVMPEDAEFTGEGHNEKRNPDADKPMCFIKRYGGFKTACKHKNQDVSAC SMOVAIPEESPTERTKLPQDPSCKRTIARRTKIRYGGFKTACKHKNQDVSAC LPLTTSYMPQGHGCGALHPQVWLTAAHCTDITKTHLKYVGLDGLDKKEFHQSTRV ERTFYSHYNEDEIPNDIALILKLPVGHGALSKLVKVCLEPDSFSPSGECHIS GGVVETGKSGRDLDAKRLIANTLCNSROLYDHTMIDSMICGNLQKPGDTCOD SSGPLTECKEDTGYVYGIVSMGLECKRPGVYTVTKLNMIKATIKSESG"

BASE COUNT 643 a 651 c 578 g 536 t

ORIGIN

Query Match 99.8%; Score 1679.8; DB 9; Length 2408;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1681; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGTTTGCCAGAGATGTCATCTCATGTCCTGTTAATGCTGTGGTGGAAAGACA 60
|||||
97 ATGTTTGCCAGAGATGTCATCTCATGTCCTGTTAATGCTGTGGTGGAAAGACA 156
|||||
61 GCCTGTGGGTTCCCTGATGTCCTTATTTGAAAGCCGGAGCCAGACGACCCCTGAC 120
|||||
157 GCCTGTGGGTTCCCTGATGTCCTTATTTGAAAGCCGGAGCCAGACGACCCCTGAC 216
|||||
121 CAGTATGATTACAGCTACGAGATTATTAATCAGAGAAGAACACAGTAGACACTTACC 180
|||||
217 CAGTATGATTACAGCTACGAGATTATTAATCAGAGAAGAACACAGTAGACACTTACC 276
|||||
181 CATGCTGAGATCTGACGTGCTACACTGAGACACCAAGTGTCCATGCCAGCCCAAC 240
|||||
277 CATGCTGAGATCTGACGTGCTACACTGAGACACCAAGTGTGATTCATGCCAGCCCAAC 336
|||||
241 CCCTGTGAACACGGTGGGACTGCTGTCATCGTGGAGACACTTACATGACAGCTGCTG 300
|||||
337 CCCTGTGAACACGGTGGGACTGCTGTCATCGTGGAGACACTTACATGACAGCTGCTG 396
|||||
301 GCTGCTTTCTGGGATTAAGTGTGCAAAAGTGAAGTACAGTCAAGAGCAACCCATGT 360
|||||
397 GCTGCTTTCTGGGATTAAGTGTGCAAAAGTGAAGTACAGTCAAGAGCAACCCATGT 456
|||||
361 GGCAGGGGCCAATGCTCATTAACAGAGTCTCCCTACTACGCTGTGTCTGTAACAC 420
|||||
457 GGCAGGGGCCAATGCTCATTAACAGAGTCTCCCTACTACGCTGTGTCTGTAACAC 516
|||||
421 CCTTACACAGTCCAGTGTCTCCCAAGTGTCTCCTGATGACAGCAACCCCTGACAG 480
|||||
517 CCTTACACAGTCCAGTGTCTCCCAAGTGTCTCCTGATGACAGCAACCCCTGACAG 576

QY 481 AATGGGGCTACTGCTCTCCGGCATTAAGCGAGATCCAAAGTTCACTGCTGCTGCCGAC 540
|||||
DB 577 AATGGGGCTACTGCTCTCCGGCATTAAGCGAGATCCAAAGTTCACTGCTGCTGCCGAC 636
|||||
QY 541 CAGTTTCAGAGGGAATTTCTGTGAAATAGAGTTCTGATGATGCTGATTTGGCATGGCTAC 600
|||||
DB 637 CAGTTTCAGAGGGAATTTCTGTGAAATAGAGTTCTGATGATGCTGATTTGGCATGGCTAC 696
|||||
QY 601 TCTTACCGAGGGAATTAATAGACAGTCAACACGATGCTGCTCTTACTGAGACACC 660
|||||
DB 697 TCTTACCGAGGGAATTAATAGACAGTCAACACGATGCTGCTCTTACTGAGACACC 756
|||||
QY 661 CACTCTCTCTGCGAGAGATTACACATGTTTATGAGAGATGCTGAACCCATGGAGTT 720
|||||
DB 757 CACTCTCTCTGCGAGAGATTACACATGTTTATGAGAGATGCTGAACCCATGGAGTT 816
|||||
QY 721 GGGAGAACAAATTTCTGAGAAACCCAGATGGGAGCGAAAGCCCTGGTGGCTTTATTA 780
|||||
DB 817 GGGAGAACAAATTTCTGAGAAACCCAGATGGGAGCGAAAGCCCTGGTGGCTTTATTA 876
|||||
QY 781 GTTACCAATGACAAAGTGAATGGAATGGAATGATGATGCTGACGCTCTGACCCAGGAC 840
|||||
DB 877 GTTACCAATGACAAAGTGAATGGAATGGAATGATGATGCTGACGCTCTGACCCAGGAC 936
|||||
QY 841 GTTGCTTACCCAGAGAAAGCCCACTGAGCCATCAACCAAGCTTCCGGGGTTTGACTCC 900
|||||
DB 937 GTTGCTTACCCAGAGAAAGCCCACTGAGCCATCAACCAAGCTTCCGGGGTTTGACTCC 996
|||||
QY 901 TGTGAAGACGTGAGATATGACAGAGAGAAAGATCAAGAGAAATCTATGAGAGCTTTAAGAC 960
|||||
DB 997 TGTGAAGACGTGAGATATGACAGAGAGAAAGATCAAGAGAAATCTATGAGAGCTTTAAGAC 1056
|||||
QY 961 ACGCGGGGCAAGCAACCCATGAGAGGGTCCCTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
|||||
DB 1057 ACGCGGGGCAAGCAACCCATGAGAGGGTCCCTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1116
|||||
QY 1021 ATGCCCCAGGGCACTTCTGTGTTGGGGCGCTGATTCACCCCTGCTGGTGTCTACTGCT 1080
|||||
DB 1117 ATGCCCCAGGGCACTTCTGTGTTGGGGCGCTGATTCACCCCTGCTGGTGTCTACTGCT 1176
|||||
QY 1081 GCCCATGACCCAGATTAACCAACAGACATCTTAAAGTGTGCTGAGGGACCGAGACTG 1140
|||||
DB 1177 GCCCATGACCCAGATTAACCAACAGACATCTTAAAGTGTGCTGAGGGACCGAGACTG 1236
|||||
QY 1141 AAGAAAGAAATTTTCATGAGCAGAGCTTTAGGTGCGAAGAGATTTCAAGTACAGCCAC 1200
|||||
DB 1237 AAGAAAGAAATTTTCATGAGCAGAGCTTTAGGTGCGAAGAGATTTCAAGTACAGCCAC 1296
|||||
QY 1201 TACAATGAAGAGATGAGATTTCCCAATGATATTGCAATTGCTCAAGTTAAAGCCAGTG 1260
|||||
DB 1297 TACAATGAAGAGATGAGATTTCCCAATGATATTGCAATTGCTCAAGTTAAAGCCAGTG 1356
|||||
QY 1261 GATGTCCTGCTGTAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 1320
|||||
DB 1357 GATGTCCTGCTGTAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 1416
|||||
QY 1321 TTTTCCCTGAGAGAGTGGAGTGGCAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
|||||
DB 1417 TTTTCCCTGAGAGAGTGGAGTGGCAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1476
|||||
QY 1381 TCCCGCCAGCTCTGAGATCCAAAGTCAAGCTGATTTGACACTTGTGCACTCCCG 1440
|||||
DB 1477 TCCCGCCAGCTCTGAGATCCAAAGTCAAGCTGATTTGACACTTGTGCACTCCCG 1536
|||||
QY 1441 CACTCTATGACACATGATGTTGATGACAGTATGATGTGCGAGAAATCTTCAGAAACCT 1500
|||||
DB 1537 CACTCTATGACACATGATGTTGATGACAGTATGATGTGCGAGAAATCTTCAGAAACCT 1596
|||||
QY 1501 GGGCAGACACCTGCCAGGCTGACTCTGAGAGCCCTTCACTGTGAGAAAGAGAGGAC 1560
|||||
DB 1597 GGGCAGACACCTGCCAGGCTGACTCTGAGAGCCCTTCACTGTGAGAAAGAGAGGAC 1656
|||||

QY 1561 TACTACGCTATGAGATAGTGGGCGCTGGAGTGTGAGAGAGGCCAGGGGCTTAC 1620
|||||
Db 1657 TACTAGCTATGGATAGTAGTGGGGCGCTGGAGTGTGAGAGAGGCCAGGGGCTTAC 1716
QY 1621 ACCAAGTTACCAAAATCTCTGATTTGGATCAAAAGCCATCAAAAGTGAAGTGGCTTC 1680
|||||
Db 1717 ACCAAGTTACCAAAATCTCTGATTTGGATCAAAAGCCATCAAAAGTGAAGTGGCTTC 1776
QY 1681 TAA 1683
|||
Db 1777 TAA 1779
RESULT 4
LOCUS AX409021 3008 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 1668 from Patent WO0229103.
ACCESSION AX409021
VERSION AX409021.1 GI:21441726
KEYWORDS
SOURCE human;
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 1668 11-APR-2002;
GENE LOGIC INC (US)
FEATURES
Source 1..3008
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. DA9742"
BASE COUNT 791 a 837 c 703 g 677 t
ORIGIN
Query Match 99.8%; Score 1679.8; DB 6; Length 3008;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1661; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGTTTGCCAGAGATGTCATCTCCATGTTCTGCTGTTAATGAGCTGTGGGAAAGACA 60
|||||
Db 97 ATGTTTGCCAGAGATGTCATCTCCATGTTCTGCTGTTAATGAGCTGTGGGAAAGACA 156
QY 61 GCCTGGGTTCTCCCGATGTCCTTTATTTGGAAGCCCTGGAGCCAGACTGGACCCCTGAC 120
|||||
Db 157 GCCTGGGTTCTCCCGATGTCCTTTATTTGGAAGCCCTGGAGCCAGACTGGACCCCTGAC 216
QY 121 CAGTATGATTACAGCTACGAGAGATTTAATCAGAGAAGAACACCACTAGACACTTACC 180
|||||
Db 217 CAGTATGATTACAGCTACGAGAGATTTAATCAGAGAAGAACACCACTAGACACTTACC 276
QY 181 CAGTATGAGATCTGAGCTGCTACTACGAGAGAACCAAGTGCATGCCAGCCCAAC 240
|||||
Db 277 CAGTATGAGATCTGAGCTGCTACTACGAGAGAACCAAGTGCATGCCAGCCCAAC 336
QY 241 CCTGTGAACAGGTGGGAGCTGCTGTCCATGAGAGACACCTTACATCAGACTGCTG 300
|||||
Db 337 CCTGTGAACAGGTGGGAGCTGCTGTCCATGAGAGACACCTTACATCAGACTGCTG 396
QY 301 GCTCCTTCTCTGGGAATTAAGTGTCAAGAAAGTCAAAATAGTGCAGAGGCAACCCATGT 360
|||||
Db 397 GCTCCTTCTCTGGGAATTAAGTGTCAAGAAAGTCAAAATAGTGCAGAGGCAACCCATGT 456
QY 361 GGGCGGGGCAATGTCTCATATTACCCAGAGTCCCTCCTACTACGCTGTGTCTTAAACAC 420
|||||
Db 457 GGGCGGGGCAATGTCTCATATTACCCAGAGTCCCTCCTACTACGCTGTGTCTTAAACAC 516
QY 421 CCTTACAGAGTCCAGCTGCTCCCAAGTGTCTCTGATGAGAGGCAAAACCCCTGCCAG 480
|||||
Db 517 CCTTACAGAGTCCAGCTGCTCCCAAGTGTCTCTGATGAGAGGCAAAACCCCTGCCAG 576

QY 481 AATGGGCTACCTGCTCCCGCATTAAGCGAATCCAAAGTTACACTGTGCTGTCCGAC 540
|||||
Db 577 AATGGGCTACCTGCTCCCGCATTAAGCGAATCCAAAGTTACACTGTGCTGTCCGAC 636
QY 541 CAGTTCAAGGGGAAATTTCTGAAATAGTTCATGATGACTCTTATGTTGGCGATGGCTAC 600
|||||
Db 637 CAGTTCAAGGGGAAATTTCTGAAATAGTTCATGATGACTCTTATGTTGGCGATGGCTAC 696
QY 601 TCTTACGAGGGAATTAATGATPAGACATCAACAGCATGCTGCTTTACTGGAATCC 660
|||||
Db 697 TCTTACGAGGGAATTAATGATPAGACATCAACAGCATGCTGCTTTACTGGAATCC 756
QY 661 CACCTCTCTTGCAGAGAAATTAACATGTTTATGAGATGCTGGAACCCATGGGAT 720
|||||
Db 757 CACCTCTCTTGCAGAGAAATTAACATGTTTATGAGATGCTGGAACCCATGGGAT 816
QY 721 GGGGAACCAATTTTCTGACAAACCCAGATGCGGAGCAAAAGCCCTGTGCTTTATTA 780
|||||
Db 817 GGGGAACCAATTTTCTGACAAACCCAGATGCGGAGCAAAAGCCCTGTGCTTTATTA 876
QY 781 GTTACCAATGACAAGTGAATGGAATGGAATGATGATGCTGAGCTGTGAGCCAGAGAC 840
|||||
Db 877 GTTACCAATGACAAGTGAATGGAATGGAATGATGATGCTGAGCTGTGAGCCAGAGAC 936
QY 841 GTTGCCTACCGAGAGAAAGCCCACTGAGCATACCAAGCTTCCGGGTTTGACTCC 900
|||||
Db 937 GTTGCCTACCGAGAGAAAGCCCACTGAGCATACCAAGCTTCCGGGTTTGACTCC 996
QY 901 TGTGGAAGACTGAGATGACGAGAGAGAGATCAAGAGATTTATGAGGCTTTTAAAGC 960
|||||
Db 997 TGTGGAAGACTGAGATGACGAGAGAGAGATCAAGAGATTTATGAGGCTTTTAAAGC 1056
QY 961 ACGGGGGCAAGCAAGCCATGAGGAGGAGGCTCCCAAGTCTGCTGCTGACATCTCC 1020
|||||
Db 1057 ACGGGGGCAAGCAAGCCATGAGGAGGAGGCTCCCAAGTCTGCTGCTGACATCTCC 1116
QY 1021 ATGCCCGAGGGCACTTGTGTGTGGGGCGCTGATCCACCCTGCTGGGTGCTCACTGCT 1080
|||||
Db 1117 ATGCCCGAGGGCACTTGTGTGTGGGGCGCTGATCCACCCTGCTGGGTGCTCACTGCT 1176
QY 1081 GCCCATGCAACCGCATTAATAAACCGACATCTTAAAGTGGTGTAGGGAGCCAGACTG 1140
|||||
Db 1177 GCCCATGCAACCGCATTAATAAACCGACATCTTAAAGTGGTGTAGGGAGCCAGACTG 1236
QY 1141 AAGAAAGAAATTTTCATGAGCAGAGCTTTAGGGTGCAGAAAGATATTCAAGTACAGCCAC 1200
|||||
Db 1237 AAGAAAGAAATTTTCATGAGCAGAGCTTTAGGGTGCAGAAAGATATTCAAGTACAGCCAC 1296
QY 1201 TACATGAAGAGATGAGATTCCTCCCAATGATATTGCAATTGCTCAAGTTAAAGCCAGTG 1260
|||||
Db 1297 TACATGAAGAGATGAGATTCCTCCCAATGATATTGCAATTGCTCAAGTTAAAGCCAGTG 1356
QY 1261 GATGGTACGTGCTCTAGAAATCAAAATAGCTGAAGACTGTGCTTGGCTGATGGGTCC 1320
|||||
Db 1357 GATGGTACGTGCTCTAGAAATCAAAATAGCTGAAGACTGTGCTTGGCTGATGGGTCC 1416
QY 1321 TTTCCCTCTGGAGTGAATGACATCTGCTGGGGGTGTTACAGAAACAGGAAAGGG 1380
|||||
Db 1417 TTTCCCTCTGGAGTGAATGACATCTGCTGGGGGTGTTACAGAAACAGGAAAGGG 1476
QY 1381 TCCCGCCAGCTCTGATGCCCAAGTCAAGCTGATGCGCAACCTTTGTGCAACTCCGC 1440
|||||
Db 1477 TCCCGCCAGCTCTGATGCCCAAGTCAAGCTGATGCGCAACCTTTGTGCAACTCCGC 1536
QY 1441 CAACTATAGACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
|||||
Db 1537 CAACTATAGACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1596
QY 1501 GGGCAAGACACCTGCCAGGATGATCTGTGAGGCCCCCTGACCTGTGAGAGAGCGGACCC 1560
|||||
Db 1597 GGGCAAGACACCTGCCAGGATGATCTGTGAGGCCCCCTGACCTGTGAGAGAGCGGACCC 1656
QY 1561 TACTACGCTATGAGATAGTAGTGGGCGCTGGAGTGTGAGAGAGGCCAGGGGCTTAC 1620

Db	1657	TACTACGCTATGGGATAGTGAAGCTGGGCGCTGGAGTGTGGAAAGAGGCGAGGGGCTAC	1716
Qy	1621	ACCCAGTTACCAAAATTCCTGAATTGGATCAACGCCACCATCAAAAGTGAAAGTGGCTTC	1680
Db	1717	ACCCAGTTACCAAAATTCCTGAATTGGATCAACGCCACCATCAAAAGTGAAAGTGGCTTC	1776
Qy	1681	TAA	1683
Db	1777	TAA	1779

Db	1357	GATGGTCACCTGCTCTAGAAATCCAAATACGTGAAGAACTGTGTGCTTGGCTGATGGCTCC	1416
Qy	1321	TTTCCCTCTGGAGAGTGAAGTGCACATCTCTGGCTGGGTGTTCACAGAAACAGAGAAAGG	1380
Db	1417	TTTTCCCTCTGGAGAGTGAAGTGCACATCTCTGGCTGGGTGTTCACAGAAACAGAGAAAGG	1476
Qy	1381	TCCGGCCACTCTGSGATGGCCAAAGTCAGCTGATGGCAACACTTTGTGCAACTTCCCGC	1440
Db	1477	TCCGGCCACTCTGSGATGGCCAAAGTCAGCTGATGGCAACACTTTGTGCAACTTCCCGC	1536
Qy	1441	CAACTCTATGACACATGATTTGATGACAGTATGATCTGTGACAGAAATCTTCAGAAACT	1500
Db	1537	CAACTCTATGACACATGATTTGATGACAGTATGATCTGTGACAGAAATCTTCAGAAACT	1596
Qy	1501	GGGCAAGACACTGCGACGGGTGACTCTGAGAGGCCCTTCGACTGTGAGAAAGCAGGCACC	1560
Db	1597	GGGCAAGACACTGCGACGGGTGACTCTGAGAGGCCCTTCGACTGTGAGAAAGCAGGCACC	1656
Qy	1561	TACTACGTCTATGGGATAGTGAGCTGGGGCTTGGAGTGTGAGAAAGGCCAGGGGTCTAC	1620
Db	1657	TACTACGTCTATGGGATAGTGAGCTGGGGCTTGGAGTGTGAGAAAGGCCAGGGGTCTAC	1716
Qy	1621	ACCCAGTTACCAAAATTCCTGAATTTGGATCAACAAGCCACATCAAAAGTGAAAGTGGCTTC	1680
Db	1717	ACCCAGTTACCAAAATTCCTGAATTTGGATCAACAAGCCACATCAAAAGTGAAAGTGGCTTC	1776
Qy	1681	TAA 1683	
Db	1777	TAA 1779	

RESULT 6	
LOCUS	BC031412
DEFINITION	BC031412 2251 bp mRNA linear PRI 27-JUN-2002
ACCESSION	Homo sapiens, hyaluronan binding protein 2, clone MGC:34485
VERSION	IMAGE:5187066, mRNA, complete cds.
KEYWORDS	BC031412 BC031412.1 GI:21618648
SOURCE	MGC.
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
JOURNAL	Strausberg, R. Direct Submission Submitted (06-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgabs-remail.nih.gov Tissue Procurement: Life Technologies, Inc. CDNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: The I.M.A.G.E. Consortium (LNL) Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/ Contact: amge@bcm.tmc.edu Gunaratne, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.
REMARK	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov Series: IRAK Plate: 50 Row: 9 Column: 21 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 20302151. Location/Qualifiers 1..2251 /organism="Homo sapiens"
FEATURES	
SOURCE	

	Query Match	Best Local Similarity	Matches 1680;	Conservative	99.7%;	Score 1678.2;	DB 9;	Length 2251;
					99.8%;	Pred. No. 0;		
						Mismatches 0;	Indels 3;	Gaps 0;
QY	1	ATGTTTGCCAGGATGTCGATCCATGTCCTGCTGTTAAATGGCTCTGTGGGAAGACA	60					
Db	81	ATGTTTGCCAGGATGTCGATCCATGTCCTGCTGTTAAATGGCTCTGTGGGAAGACA	140					
QY	61	GCCTGTGGGCTCTCCCTGATGTCCTTATTTGGAAAGCCTGTGACCCAGATGGACCCCTGCAC	120					
Db	141	GCCTGTGGGCTCTCCCTGATGTCCTTATTTGGAAAGCCTGTGACCCAGATGGACCCCTGCAC	200					
QY	121	CAGTATGATTTACACTACGAGGATTAATAGAGAAGAACCCAGTAGACACTTTACC	180					
Db	201	CAGTATGATTTACACTACGAGGATTAATAGAGAAGAACCCAGTAGACACTTTACC	260					
QY	181	CATGCTGAAGAACTCTGACTGTAACACTAGACACCAAGCTGATCCCTGCGAAGCCACAC	240					
Db	261	CACGCTGAAGAACTCTGACTGTAACACTAGACACCAAGCTGATCCCTGCGAAGCCACAC	320					
QY	241	CCCTGTGAACACGGTGGGGAACGCTCGCTCCATGAGGACACCTTACATGACGCTGCG	300					
Db	321	CCCTGTGAACACGGTGGGGAACGCTCGCTCCATGAGGACACCTTACATGACGCTGCG	380					
QY	301	GCTCCTTTCTCTGGGAATAGTGTGAGAAAGTGCAGAAATACCTGCAGAGCAACCCATGT	360					
Db	381	GCTCCTTTCTCTGGGAATAGTGTGAGAAAGTGCAGAAATACCTGCAGAGCAACCCATGT	440					
QY	361	GGCGGGGGGCAATGTCATTTACCCAGAGGTCCTCCCTACTACCGTGTGTCTAAACAC	420					
Db	441	GGCGGGGGGCAATGTCATTTACCCAGAGGTCCTCCCTACTACCGTGTGTCTAAACAC	500					
QY	421	CCTTACACAGTCCCAAGCTGCTCCCAAGTGTCTGTATGACGAGGCCAAGCCCTCCGAC	480					
Db	501	CCTTACACAGTCCCAAGCTGCTCCCAAGTGTCTGTATGACGAGGCCAAGCCCTCCGAC	560					
QY	481	AATGGGGGTACCTCTCCCGGATTAAGGGGAGATCAATTACCTGCTGTCTCCGAC	540					
Db	561	AATGGGGGTACCTCTCCCGGATTAAGGGGAGATCAATTACCTGCTGTCTCCGAC	620					
QY	541	CAGTTTCAAGGGAATTTCTGTGAAATAGTTTGTGATGACTGCTATGTTGGCGATGGCTAC	600					
Db	621	CAGTTTCAAGGGAATTTCTGTGAAATAGTTTGTGATGACTGCTATGTTGGCGATGGCTAC	680					
QY	601	TCTTACCCAGGGAATTAATGAGTACAGTCAACACATGAGGTGCTTTACTGTGAAATCC	660					
Db	681	TCTTACCCAGGGAATTAATGAGTACAGTCAACACATGAGGTGCTTTACTGTGAAATCC	740					

QY 661 CACCTCTCTTGCAGAGATTACAAATGTTTATGAGAGATGCTGAAACCATGGGATT 720
 |||||
 DB 741 CACCTCTCTTGCAGAGATTACAAATGTTTATGAGAGATGCTGAAACCATGGGATT 800
 QY 721 GGGGAACACATTTTCTGAGAAACCAAGTGGGAGGAAAGCCCTGGCTTTATTTAA 780
 |||||
 DB 801 GGGGAACACATTTTCTGAGAAACCAAGTGGGAGGAAAGCCCTGGCTTTATTTAA 860
 QY 781 GTTACCAATGACAAAGTGAATGGGAATCTGTAGTCTGAGCTGCTGAGCCAGAGC 840
 |||||
 DB 861 GTTACCAATGACAAAGTGAATGGGAATCTGTAGTCTGAGCTGCTGAGCCAGAGC 920
 QY 841 GTTGTCTACCCAGAGGAAAGCCCATGAGCCATCAACCAAGCTCCGGGGTTTGCTCC 900
 |||||
 DB 921 GTTGTCTACCCAGAGGAAAGCCCATGAGCCATCAACCAAGCTCCGGGGTTTGCTCC 980
 QY 901 TGTGGAAGACTGAGATAGCAGAGAGGAAGATCAAGAGATCTATGAGGCTTTAAGAGC 960
 |||||
 DB 981 TGTGGAAGACTGAGATAGCAGAGAGGAAGATCAAGAGATCTATGAGGCTTTAAGAGC 1040
 QY 961 AGGGGGGCAAGCACCATGGAGGGGCTCTGAGTCTGCTGCTGACCATCTCC 1020
 |||||
 DB 1041 AGGGGGGCAAGCACCATGGAGGGGCTCTGAGTCTGCTGCTGACCATCTCC 1100
 QY 1021 ATGCCCCAGGGCCACTTCTGTGTGGGGCGCTGATCCACCCCTCTGGTCTCACTGCT 1080
 |||||
 DB 1101 ATGCCCCAGGGCCACTTCTGTGTGGGGCGCTGATCCACCCCTCTGGTCTCACTGCT 1160
 QY 1081 GCCCACTGACGACATAAAAACACAGATCTAAAGTGGTGGTAGGGAGCCAGAGCTG 1140
 |||||
 DB 1161 GCCCACTGACGACATAAAAACACAGATCTAAAGTGGTGGTAGGGAGCCAGAGCTG 1220
 QY 1141 AAGAAAGAAATTTTCATGAGCAGAGCTTTAGGGTGCAGAAAGATTTCAAGTACACCC 1200
 |||||
 DB 1221 AAGAAAGAAATTTTCATGAGCAGAGCTTTAGGGTGCAGAAAGATTTCAAGTACACCC 1280
 QY 1201 TACATGAAGAGATGAGATTCCTCCCAATGATTTGCTCAAGTTAAAGCCAGTG 1260
 |||||
 DB 1281 TACATGAAGAGATGAGATTCCTCCCAATGATTTGCTCAAGTTAAAGCCAGTG 1340
 QY 1261 GATGTCACCTGTGCTTGAATCCAAATACGAGAGACCTGTGCTGCTGATGGGCTC 1320
 |||||
 DB 1341 GATGTCACCTGTGCTTGAATCCAAATACGAGAGACCTGTGCTGCTGATGGGCTC 1400
 QY 1321 TTTCCCTGGGAGTGAAGTGCACATCTCTGCTGGGGTGTACAGAAACAGGAAAGG 1380
 |||||
 DB 1401 TTTCCCTGGGAGTGAAGTGCACATCTCTGCTGGGGTGTACAGAAACAGGAAAGG 1460
 QY 1381 TCCCGGACACTGCTGGATGCCAAGTCAAGCTGATGGCAACACTTTGTGCACTCCGC 1440
 |||||
 DB 1461 TCCCGGACACTGCTGGATGCCAAGTCAAGCTGATGGCAACACTTTGTGCACTCCGC 1520
 QY 1441 CAACCTATGACACATGATGATGACAGTATGATCTGTGAGAAATCTTCAGAAACCT 1500
 |||||
 DB 1521 CAACCTATGACACATGATGATGATGACAGTATGATCTGTGAGAAATCTTCAGAAACCT 1580
 QY 1501 GGGCAAGAACTGCGCAGAGTGAAGTGGAGGCCCCCTGACCTGTGGAAGAGAGGAGCC 1560
 |||||
 DB 1581 GGGCAAGAACTGCGCAGAGTGAAGTGGAGGCCCCCTGACCTGTGGAAGAGAGGAGCC 1640
 QY 1561 TACTAGCTATGAGATGAGTGGGAGCTGGGAGCTGAGTGTGAGAGAGGCGGAGGCTTAC 1620
 |||||
 DB 1641 TACTAGCTATGAGATGAGTGGGAGCTGGGAGCTGAGTGTGAGAGAGGCGGAGGCTTAC 1700
 QY 1621 ACCCAGTTACCAATTCCTGAATTTGATCAAGCCACATCAAAAGTGAAGAGGCTTC 1680
 |||||
 DB 1701 ACCCAGTTACCAATTCCTGAATTTGATCAAGCCACATCAAAAGTGAAGAGGCTTC 1760
 QY 1681 TAA 1683
 |||||
 DB 1761 TAA 1763

RESULT 7
 BC031775
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REMARK
 COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HS3C
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Huiyk, S.M., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRMA Plate: 37 Row: n Column: 14
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.
 Location/Qualifiers

FEATURES
 source
 1. 2119
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="FVB/N"
 /clone="MGC:28705 IMAGE:424257"
 /tissue_type="Kidney, normal, 5 month old male mouse."
 /clone_lib="NCI_CGAP_Kid14"
 /lab_host="DH10B"
 /note="Vector: pCMV-Sport6"
 59..1612
 /codon_start=1
 /product="unknown (protein for MGC:28705)"
 /protein_id="AAH31775.1"
 /db_xref="GI:21594555"
 /translation="MLFVRLLLALVGRSVIGLSLMSFLAPDPDPCGNSPCEHGGD
 CTRIDTFSCSPAPFSGSRCTAONKCKDNPCVGDCLITOKHPYRACRYPTGP
 DCSKVLPAICRNPONGCVRSHRRRSRFTCAPOYXGKFCLEIGDDCYVDGYSYR
 GKVSIVNONPCLVWNSHLLLOFTYWMEDAEETHGLAENHPCRNPDGDKPMCEYVK
 NSEKIKMEYCVYVCPYPTNPVSESLLEPWEVLPGEFCCKTEVAEHVAKRLVGGFK
 STRGNHMOVSLQSLPLTTSMPOGIFCGGALIHPCWLTAAHCHDILKRIHKVYIGD
 QDKRTESHEDTFVEKILVTSYNERDELPNDILALKLPVGGHCALESRYVTVY
 LPSDFPFGTICHTISGWSVETEGSRQLDKNVLLANPLCNRSROLVHFTIDDSMIC
 AGNLQKPGSDPCGDSGSGPLTCEKDDTYVVGIVSMGQECGRKGVYTVTFLWMIK
 TMHREAGL"
 CDS

BASE COUNT 577 a 561 c 530 g 451 t
 ORIGIN

Query Match 62.2%; Score 1047; DB 10; Length 2119;
 Best Local Similarity 82.2%; Pred. No. 2,5e-305;
 Matches 1203; Conservative 0; Mismatches 260; Indels 0; Gaps 0;
 QY 221 CTGATCATGCGACCCGCTGTGAACAGCGTGGGAGACTCCTCTGCTCATGGAGCA 280

```

Db      150  CAGATCCATGGCAGTCCAAACCCCTGTGAACACGGGGGAGCTGATCATCAGAGGGAGTA 209
Oy      281  CTTTACATGAGTGGCTGGCTCTTCTCTGGGAATAGCTCGAGAAAGTGCATAATA 340
Db      210  CCTTCAGTTGAGTGGAGCCGCCCTTCTGGGGAGCGGTGGAGCTGCAGAAACA 269
Oy      341  CCGTGAAGACAAACCCATGGCGGGGCAATGCTCATATACCAGAGTCCCTACT 400
Db      270  AGTGAAGGACAAACCATGTCTCATGTGATGCTCATATACCAGAGAACCCCTACT 329
Oy      401  ACCGTGTGCTGTAAACACCCCTACAGAGTCCAGCTGCTCCCAAGTGGTCTGTAT 460
Db      330  ACCGTGTGCTGTAAACACCCCTACAGAGTCCAGCTGCTCCCAAGTGGTCTGTAT 389
Oy      461  GCAGGCAAAACCCCTGCCAGAAATGGGGCTACCTGCTCCGGCAATAGCGAGATCCAAGT 520
Db      390  GCAGGCAAAACCCCTGCCAGAAATGGGGAGTCTGTCCGACACAGAGAGATCCAGGT 449
Oy      521  TCACCTGTGCTGTCCCGAGAGTTCAGGGGAAATTTCTGTAAATAGTTCGTACT 580
Db      450  TTACCTGTGCTGTCCCGAGAGTTCAGGGGAAATTTCTGTAAATAGTTCGTACT 509
Oy      561  CCTATGTTGGCGATGCTACTCTTACGAGGGAAATGAATAGGACATCAACGACATG 640
Db      510  GTTATGTCGGTGGTACTCTTACCGAGGCAAAAGTGAAGACATGACACGACGAAAC 569
Oy      641  CGTGCTTTTACTGGAACCTCCGACCTCTCTGAGGAGGAATTCACATGTTATGAGG 700
Db      570  CATGCTTTTACTGGAACCTCCGACCTCTCTGAGGAGGAATTCACATGTTATGAGG 629
Oy      701  ATGCTGAACCCATGGGATGGGGAAACAATTTCTGCAGAAACCCAGATGGGAGCAA 760
Db      630  ATGAGAGACCCAGGAGTGCAGAGACCAACTCTGCAGAAACCCAGATGGAGCCACA 689
Oy      761  AGCCCTGTGCTTTTAAAGTACCAATGACAGGAAATGGGAATGCTGTGCT 820
Db      690  AACCTGTGTTTCTGCAGAGTGAACAGTGAAGAGTGAAGTGGAAATGCTGTGCTGA 749
Oy      821  CAGCTCTTACGCCAGACGTTGCTTACCCAGAGAAACCCCACTGAGCCATCAACA 880
Db      750  CAGTCTTCCAGAGCTGCACCCCTTAACCCAGTGAAGAACCTTGTGAGCCTGTGATG 809
Oy      881  AGCTTCCGGGTTTACTCTGTGGAAGACTGATGATGAGAGGAATCAAGAA 940
Db      810  AGTCCAGGAGTTCAGTCTCTCGGGAAGAGAGAGTACTGAAACCGACATGAAGGTA 869
Oy      941  TCTATGAGAGCTTTAAGACAGCGGGGCAAGCAACCATGAGCGCTCCCTCCAGCT 1000
Db      870  TCTACGGGGGCTTTAAGACAGAGAGCAAGCAACCGTGGCAGGTGCCGACGACT 929
Oy      1001  CGCTGCTTGAACATCTGCATGCCCCAGGGGCACTTGTGGTGGGGCCGATCCACC 1060
Db      930  CACTGCGGTTGACACCTTCATGCCCAAGGCCACTTCTGTGGGGCCCTCATCCACC 989
Oy      1061  CCGTGGGGGCTGCTGCTGCTGCCACATGCAATATAAACCAAGACATCTAAAGGTG 1120
Db      990  CTTGCTGGGGGCTGCTGCTGCTGCCACATGCAATATAAACCAAGACATCTAAAGGTG 1049
Oy      1121  TGTAGGGGACAGGACCTGGAAGAAAGAAATTTATGACGAGAGCTTTAGGGTGA 1180
Db      1050  TACTAGGGGATCAGGACCTGAAGAAAGACAGATCCATGACGAGCTTCAAGGTGAAA 1109
Oy      1181  AGATTTTCAAGACAGCAGTACAAATGAAGAGATGATTCCTCCCAATATATTCAT 1240
Db      1110  AATACTGAAGTACAGTACATGATTAAGAAAGATGATTCCTCCCAATATATTCAT 1169
Oy      1241  TGTCAAGTTAAAGCAGTGTGATGCTGCTAGATCCAAATACGTGAAGACTG 1300
Db      1170  TGTCAAGTTAAAGCAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1229
Oy      1301  TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1360

```

```

Db      1230  TATGTTGCCACGACCCCTTTTCCCTCTGTGAACTGAGTCCACATCTGTGGGGTG 1289
Oy      1361  TTACAGAAACAGGAAAGAGGTCCCGCAGCTCTGTGATGCCAAATGATGATGATGCA 1420
Db      1290  TTACAGAAACAGGAAAGAGGTCCCGCAGCTCTGTGATGCTAAATGATGATGATGCA 1349
Oy      1421  ACAATTTGTGCAACTCCCGCAACTATGACACATGATTAATGATGATGATGATG 1480
Db      1350  ACCCTTTGTGCAACTCCCGCAACTATGACACATGATTAATGATGATGATGATG 1409
Oy      1481  CAGAAATTTTGAAGACCTGGGCAAGACACCTGCCAGGGTACTCTGTGAGGCCCTGA 1540
Db      1410  CGGGGAACCTTCAGAAAGCCCGATCAGACACCTGCCAGGGTACTCTGTGAGGCCCTGA 1469
Oy      1541  CTTGTGAGAGAGCGGACCTTACTGATGATGATGATGATGATGATGATGATGATG 1600
Db      1470  CTTGTGAGAGAGATGAACTTACTGATGATGATGATGATGATGATGATGATGATG 1529
Oy      1601  AGAAGAGCCAGGGGCTTACACCCAGTTCACAAATTCCTGAATGATGATGATGATG 1660
Db      1530  GGAAGAGCCAGGAGCTTACACTCAAGTCAACGATTCCTGAATGATGATGATGATG 1589
Oy      1661  TCAAGTGAAGTGGCTTCTAA 1683
Db      1590  TGCACAGGAGGCTGCTCTGA 1612

RESULT 8
AL390197
LOCUS
DEFINITION
Human DNA sequence from clone RP11-190F19 on chromosome 10,
complete sequence.
AL390197
VERSION
AL390197.19 GI:15387800
KEYWORDS
HTG.
SOURCE
Human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 188349)
REFERENCE
Tracey, A.
direct submission
Submitted (29-AUG-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
requests: clonerequests@sanger.ac.uk
On Aug 31, 2001 this sequence version replaced gi:14596344.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats: all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; SW:
SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 10, constructed by the Sanger Centre Chromosome 10
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr10
RP11-190F19 is from the library RPI1-11.1 constructed by the group
of Pletier de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBace3.6
This sequence is the entire insert of clone RP11-190F19 The true

```

FEATURES	source	location/qualifiers
	1..188349	/organism="Homo sapiens"
		/db_xref="Laxon:9606"
		/chromosome="10"
		/clone="RP11-190F19"
		/clone_1ib="RPC1-11.1"
	22142..22187	
misc_feature	/note="Sequence from overlapping clone RP11-481H12 (A1133482). Assembly confirmed by restriction digest."	
BASE COUNT	52413 a	42671 c 41477 g 51788 t
ORIGIN		
Query Match	15.5%;	Score 261.2; DB 9; Length 188349;
Best Local Similarity	93.8%;	Pred. No. 5.1e-67;
Matches 272;	Conservative 0;	Mismatches 18; Indels 0; Gaps 0;
QY 805	GATACGTGTGATGTCCTCAGCGCTGCACGCCAGGACGCTTGCACCCAGGAGAAAGCCCC	864
Db 96119	GAAATGACCAATTATCTTTCTTGTGTGCTCCACAGACGTTCTTACCCAGAGAAAGCCCC	96178
QY 865	ACTGAGCCATCAACCAAGCTTCGCGGGTTGACTCTGTGAAAGATAGATAGCAGAG	924
Db 96179	ACGAGACCATCAACCAAGCTTCGCGGGTTGACTCTGTGAAAGATAGATAGCAGAG	96238
QY 925	AGGAAGATCAAGAAATCTATGAGAGCTTTAAGAGCAGGCGGAGACACCATGTGACAG	984
Db 96239	AGGAAGATCAAGAAATCTATGAGAGCTTTAAGAGCAGGCGGAGACACCATGTGACAG	96298
QY 985	GGCTCCCTCCAGTCCCTGCTGCTCTGACCATCTCCATGCCCCAGGCGCATCTTGCGT	1044
Db 96299	GGCTCCCTCCAGTCCCTGCTGCTCTGACCATCTCCATGCCCCAGGCGCATCTTGCGT	96358
QY 1045	GGGGCGCTGATCAACCCCTGCTGGGTGCTCACTGCCCCATGTCACCGA	1094
Db 96359	GGGGCGCTGATCAACCCCTGCTGGGTGCTCACTGCCCCATGTCACCGA	96408
RESULT 9		
AC006097/c	157750 bp	DNA linear HTG 30-AUG-2001
LOCUS	Homo sapiens chromosome 10 clone C17987SK-1031615 map 10q25, 1	
DEFINITION	Ordered piece.	
ACCESSION	AC006097.1	GI:3962489
VERSION	AC006097	
KEYWORDS	HTG; HTGS_PHASE2; HTGS_CANCELLED.	
SOURCE	Homo sapiens.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	
AUTHORS	Smith,D.R.	
JOURNAL	Sequencing of Human Chromosome 10	
REFERENCE	unpublished	
AUTHORS	Smith,D.R.	
JOURNAL	Direct Submission	
TITLE	Submitted (04-BEC-1998) Genome Therapeutics Corporation, 100 Beaver Street, Waltham, MA 02154, USA	
NOTE	Clone was sequenced in 1996 using Multiplex DNA Sequencing Technology. Data may contain low quality sequence and BAC/Cosmid vector sequences.	
COMMENT	<p>* NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces * is believed to be correct as given, however the sizes * of the gaps between them are based on estimates that have * provided by the submitter.</p> <p>* This sequence will be replaced</p> <p>* by the finished sequence as soon as it is available and * the accession number will be preserved.</p> <p>1 157750: contig of 157750 bp in length.</p>	

FEATURES		Location/Qualifiers	
source		1..157750	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/chromosome="10"	
		/map="10q25"	
		/clone="Cl1987SK-1031G15"	
BASE COUNT	43060 a 35424 c 36360 g 42895 t	11 others	
ORIGIN			
Query Match	15.3%;	Score 258;	DB 2; Length 157750;
Best local Similarity	93.1%;	Pred. No. 4.6e-66;	
Matches 270;	Conservative 0;	Mismatches 20;	Indels 0; Gaps 0;
OY	805	GAATACCTGTGATGTCACGCTGCTCAAGCCAGAGAGTTGCTTACCAGAGAAAGCCCC	864
Db	57685	GAATAGCAATTTATCTTTCTTGTTGGTCCCAACAGCTGGCTACCCAGAGAAAGCCCC	57626
OY	865	ACTGAGCCATCAACCAAGCTCCCGGGGTTTACCTCGTGGAAAGACATGATATGACAGAG	924
Db	57625	ACTGAGCCATCAACCAAGCTCCCGGGGTTTACCTCGTGGAAAGACATGATATGACAGAG	57566
OY	925	AGGAAGATCAAGAGATCTATGAGAGCTTTAAAGACAGCGCGGACCAACACCCTGGCGAG	984
Db	57565	AGGAAGATCAAGAGATCTATGAGAGCTTTAAAGACAGCGCGGACCAACACCCTGGCGAG	57506
OY	985	GGGTCCCTCCAGTCTCTGCGCTGCTGACATCTCCATATGCCGCCAGG93CACTTCTGTGT	1044
Db	57505	GGGTCCCTCCAGTCTCTGCGCTGCTGACATCTCCATATGCCGCCAGG93CACTTCTGTGT	57446
OY	1045	GGGGGGCGATGATCAAGCCGCTGGGGGTGCATCTGCTCCCATGACACCGA	1094
Db	57445	GGGGGGCGATGATCAAGCCGCTGGGGGTGCATCTGCTCCCATGACACCGA	57396
RESULT 10			
AX305372	AX305372	397 bp	DNA linear PAT 11-DEC-2001
LOCUS	Sequence 123 from Patent WO0188188.		
DEFINITION	AX305372		
ACCESSION	AX305372.1	GI:1764923	
VERSION			
KEYWORDS	house mouse.		
SOURCE	Mus musculus		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 Ishikawa, K., Asai, S., Takahashi, Y., Negata, T. and Ishii, Y.		
AUTHORS	Method for examining ischaemic conditions		
TITLE	Patent: WO 0188188-A 123 22-NOV-2001;		
JOURNAL	School Juridical Person Nihon University (JP)		
FEATURES	Location/Qualifiers		
source	1..397		
	/organism="Mus musculus"		
	/db_xref="taxon:10090"		
BASE COUNT	103 a 105 c 111 g 78 t		
ORIGIN			
Query Match	14.3%;	Score 241;	DB 6; Length 397;
Best local Similarity	84.4%;	Pred. No. 3e-61;	
Matches 271;	Conservative 0;	Mismatches 50;	Indels 0; Gaps 0;
OY	1363	ACAGAAACAGAGAAAGGATCCGCCAGCTCGTGGATGCCAAAGTCAAGCTGATGGCCAAAC	1422
Db	13	ACAGAAACAGAGAAAGGATCCGCCAGCTCGTGGATGCCAAAGTCAAGCTGATGGCCAAAC	72
OY	1423	ACTTTTGTGCACTCCCGCCCACTCTATGACACATGATATGATACAGTATGATCTGTGCA	1482
Db	73	CCTTTGTGCACTCCCGCCCACTCTATGACACATGATATGATATGATCTGTGCG	132
OY	1483	GGAATCTTCAGAAACCTGGGCAAGACCTTCCAGGGTGAATCTCTGGAGGCCCTCTGACC	1542
Db	133	GGAATCTTCAGAAACCTGGGCAAGACCTTCCAGGGTGAATCTCTGGAGGCCCTCTGACC	192


```
Db 977 GTCTGCTGGCCCTGGGCCCCCTCAGCCTTACTGCGGGAACCCAGACAGATGAGAGGCT 1036
Qy 766 TGGTGCTTTTAAATTTACCAATGACAAAGTGAAATGGAAATGCTGATCTTCACCC 825
Db 1037 TGGTGCTATGCT---GGTGAAGAGCAACCCACTGCTGTGGAGATGATTTGCGCGCTGACACC 1093
Qy 826 TGTCTACAGCCAGGACGCTTCCCTTACCCAGAGAAAGCCCACTGACGATCAACCAAGCTT 885
Db 1094 TGTGAATTCCTGGGCGAGAGTCCATCCCA-----AACCCGGAGATCCTTACACCCCTG 1147
Qy 886 CCGGGGTTTGAATCTCTGTGAAAAGCTGAGATPACAGAGAGAAATCAAGAAATCTA- 944
Db 1148 CCCCAGTACAGCCCGGCTGTGGCTCTACCTGTGCAAGAGGACAAAGAGAGGAGAGCTTC 1207
Qy 945 TGAAGGCTTTAAGAGCAGGCGGAGACACCCATGGAGCGCTCCCTGCTACCTCTCCCT 1004
Db 1208 TTGAGACACGACATCATCGGGGGGCTCATCATCTTGGCTTGCTACACCCCTGCGCTGGCT 1267
Qy 1005 GCCTCTGACCATCTCATGCCCCAGGCGCACTTCTGTGTGGGCGCTGATCCACCCCTG 1064
Db 1268 GCGATCTACATTTGGGAATA-----GCTTCTGTGGGAGGCTTGTCTCATACCTG 1317
Qy 1065 CTGGGTGCTACTGTGCTGCACTGCAACGACATTAACCAAC-----AGACATCTAAGGT 1118
Db 1318 CTGGGTGATGTGTGACAGCCCACTGCTTGGCCACAGCGCCCGCCAGGAGACAGATCAGCT 1377
Qy 1119 GGTCTAGGGGAGCAGGACCTGAAAGAAAGAAATTTGATGAGCAGAGCTTTAGGGTCA 1178
Db 1378 GGTACTGGGTACAGACTTCTTCAACCGGACAGCGAATGTGACAGACATTTGGCATTTGA 1437
Qy 1179 GAAGATATTCAGTACAGCCACATACATGAAGAGATGAGATTTCCCAATGATATTGCT 1238
Db 1438 GAAGATATGTCCTTACACCCCTGACTCGGTGTCAACCCCA---ACAACCATGACCTTGT 1494
Qy 1239 ATTTGCTCAATTAAGCAGCAGTGGATGCTGCTGCTGAAATTCATTTAGCTGAAGAC 1298
Db 1495 CTGATCTGGCTTAAAGAAAGAGAGAGCGCTGTGCTGCTCGCTCCCAATTTGTCAACC 1554
Qy 1299 TGTGTGCTGCTGCTGATG-----GGTCTTCTCCCTGAGAGTGAATGGCAGCATCTGG 1352
Db 1555 CATCTGCTGCTGAGGAGGAGGAGCGCTCTCTCCATGAGACAAATGTCAATTTGGCATG 1614
Qy 1353 CTGGGG-----TGTACAGAAACAGAAAGGTTCCCGCCAGCTCTTGATGCCAAAGT 1406
Db 1615 CTGGGGCCACATGATGAAGAAATGTGAGCAGCTACTCCACTGCTGTGAGGCACTGGT 1674
Qy 1407 CAAGCTGATTTGCCAACATTTTGTGCAACTCCCGCAACTATGACCAATGATGATGA 1466
Db 1675 CCGCTTGTGCTGACGACAGCAAGTGTAGCAGGCCAGAGATGATGCTGATGATGAGCC 1734
Qy 1467 CAGTATGATCTGTGACAGAAATCTTCAAAACCTGGGCAAGACACTGCCAGGGTGACTC 1536
Db 1735 TTAACATCTCTGTGGCGGCTACTT---GCACTGCAAACTCCGATGCTGCAAGGGGAGCTC 1791
Qy 1537 TGAAGGCCCCCTGACCTGTGAGAGAGAGGACCACTTACTAGTGTATGAGATAGTGAAGT 1586
Db 1792 AGGTGGGCGCTTGCTGTGAGAGAAATGCTGTGGCTTACTCTGTATGAGATGATGAGCTG 1851
Qy 1587 GGGGCTGAGATGAG-----AAGAGCCAGGGGCTTACACCCAAAGTTTACCAATTTCT 1640
Db 1852 GGGTGAATGCTGTGGCGGCTCAACAGCCAGAGTCTTACACCGGATGTCGCAATTAATGT 1911
Qy 1641 GAATTTGATCAAAAGCCACCAT 1661
Db 1912 GGACTGATCAAGCAGCCGAT 1932
```

```
RESULT 13
AF224724 2063 bp mRNA linear ROD 30-APR-2001
LOCUS Mus musculus hepatocyte growth factor activator HGF mRNA, complete
DEFINITION cds.
ACCESSION AF224724
```

```
VERSION AF224724.1 GI:6980071
KEYWORDS Mus musculus.
SOURCE Mus musculus.
ORGANISM Mus musculus; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2063)
AUTHORS van Adelsberg, J., Sehgal, S., Kukes, A., Brady, C., Barsch, J.,
Yang, J., and Huan, Y.
TITLE Activation of hepatocyte growth factor (HGF) by endogenous HGF
activator is required for metanephric kidney morphogenesis in vitro
JOURNAL J. Biol. Chem. 276 (18), 15099-15106 (2001)
MEDLINE 21226753
PUBMED 11032833
REFERENCE 2 (bases 1 to 2063)
AUTHORS van Adelsberg, J.S., Sehgal, S., Kukes, A., Brady, C., Barsch, J.,
Yang, J., and Huan, Y.
TITLE Direct Submission
JOURNAL Submitted (13-JAN-2000) Medicine and Anatomy and Cell Biology,
Columbia University, 630 West 168th Street, New York, NY 10032, USA
FEATURES
source
1. 2063
/oranism="Mus musculus"
/db_xref="taxon:10090"
5. 1966
/codon_start=1
/product="hepatocyte growth factor activator HGF"
/protein_id="AF34712.1"
/db_xref="GI:6980072"
/translation="MGRQAMISLCPPLPPEFLLLLLVPRGAOPQAGRNHTPEP
GPNVATPVPTIPVISGNVSTRESAPARETPOSERPPSPSSPPGGVLTBSG
OCREPPRGKMLHSCSTESASVAKWCAHTHNDPRAKCAVTLTPVSGVLTBP
CASPLCNGSTCSSTHDHSYHCSPLATFKDGTCEKDERIERYEPVEDHARRS
EGHYEQCGEGARCEDTHTACLSPLCLNGVCHLTVGTGVSVAVALGLGPHAY
IVPEHFCLENGEYRGVASTASGLSCLANSDLYLDELHVDVSAVAALVGLGPHAY
CRNDKDERPWCYVNDNALSWECRLTCAARVSHQPELALPESAPVAPPC
GKRHKRTFLPRLITGSSSLPSSHPLALYIGNSCAGSLVHVCWVSAHGFANS
PPRDSITVVGQHEFNRTDVTQGTGIEKYVYVLYSVFNPNNHDLVLRKKGERC
AVRSQFQPCLEPAGSFPPTGKCOLAGMGMENVSNSLLEALVPLVAHKS
SPEVYGDISPMLCAGFYDCSDPCGDSGSLPCEKKNVAVLVGIISMDGGRIN
KPYVTVNAVTVMDIRIRPKRVAT"
BASE COUNT 438 a 630 c 576 g 419 t
ORIGIN
Query Match 7.7%; Score 129.8; DB 10; Length 2063;
Best Local Similarity 50.9%; Pred. No. 1.8e-27;
Matches 621; Conservative 0; Mismatches 547; Indels 53; Gaps 11;
```

```

Db      1094 TGTGATCCCTGCGCCAGAGTCCACTCCCA-----AACCCCGAGATCTTAGACGCCCTTG 1147
      886 CCGGGGTTTGACTCCTCTGGAAGACAGATGACAGAGAGAGAGATCAGAGATCTA- 944
      1148 CCGGATCGACCCCGGGCTGCTGCTACTCTGTGCAAGAGGCAACAAGAGACCTTC 1207
      945 TGGAGCCTTTAAGACAGCGCGGGGAGCAGCCATGAGAGGCTCCCTCAGTCTCGCT 1004
      1208 TTGAGACACAGCATCATCGGGGGCTCATCTTTTCCCTGCTACACACCCCTGGCTGGCT 1267
      1005 GCCTGACACATCTCCATGAGCCCGGAGGAGCCTTCTGTGGGGGCTGATCCACCCCTG 1064
      1268 GGCATCTACATTTGGGAATA-----GCTTCTGTGCGGGAGCCTTGCCATCTCTG 1317
      1065 CTGGGTGCTCACTGCTGCCACTGCGACCGACATMAAAC-----AGACATCTAAAGT 1118
      1318 CTGGTAGTGTCTGACGCCCTCTTCCGCAACAGCCCGCCAGGAGACAGCATCACAGT 1377
      1119 GGTGCTAGGGGACAGGACCTGAAAGAAAGAAATTTCAAGACAGAGCTTTAGCGTCA 1178
      1378 GGTACTGGGCTGACGCTTTCAACCGCACCGACGATGTGACACAGCATTTGGCATTTGA 1437
      1179 GAAGATATTCAGTAGTACAGCCTACTCAATGAAAGATGATGATTCGCCCAATGATATTCG 1238
      1438 GAAGTATGTGCGCTTACACCCCTGTACTGGGTTCACACCCCA---ACAACCATGACTTGT 1494
      1239 ATTGCTCAAGTTAAAGCCATGATGTGCTACTGTCTTACATTCGAATTCGAAGAC 1298
      1495 CTGTGATCGCGCTTAAAGAAAGAGAGCGCGCTGTGCTGCCCTCCAGTTTGTCAACC 1554
      1299 TGTGCTCTTGGCTGATG-----GTTCTTTTCCCTCTGGAGTGTGAGTCCCATCTCTGG 1352
      1555 CATCTGCTGCTGAGGACGAGCAGCTCTTCCCTTGTGAGACACAAGTGTGAGATGGAGG 1614
      1353 CTGGGG-----TGTACAGAAACAGSAAAGGCTCCGACGCTCTCGATGCCAAAGT 1406
      1615 CTGGGGCCACATGATGTAAGTGTGACACACTTCAACTCCCTGCTGGAGGAGCTGGGT 1674
      1407 CAAGCTGATTTGGCAACACTTTGTGCACTCCGCGCAACTACTATGACACATGATGATGA 1466
      1675 CCTCTTGTGCTGACACCAAGTGTAGACACCCAGAGATATGGTGTGACATCAGGCC 1734
      1467 CATATGATCTGTGAGGAATCTTCAGAACTGGGCAAGACACTGCCAGGGTGACTC 1526
      1735 TAACTATGCTGTGCGCTTACTT---CGACTGCAAGTCCGATGCTGCGAGGGGAGCTC 1791
      1527 TGGAGCCCCCTGACTGTGTGAGAGAGGACGCACTTACTAGCTTATGGGATAGAGCTG 1586
      1792 AGGTGGGCTTGGTGTGTGAGAAATGTTGGCTTACCTGTATGGCATCATCAGCTG 1851
      1587 GGGCTGAGAGTGTAG-----AAGAGCCAGGGGCTTACACCCCAATTTACCAATTCCT 1640
      1852 GGGTATGTGGCTGTGGGCGCTCAACAGGAGAGTACACCGGTGGGCCAATTTATGT 1911
      1641 GAATTGATCAAGGCCAAT 1661
      1912 GGACTGATCAACGACGCTAT 1932

```

```

RESULT 14
BC019376      2067 bp      mRNA      linear      ROD 07-AUG-2002
LOCUS      BC019376
DEFINITION      Mus musculus, clone MGC:30305 IMAGE:5134851, mRNA, complete cds.
ACCESSION      BC019376
VERSION      BC019376.1 GI:18044172
KEYWORDS      MGC.
SOURCE      house mouse.
ORGANISM      Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 2067)
AUTHORS      Strausberg, R.

```

```

TITLE      Direct Submission
JOURNAL      Submitted (19-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK      NIH-MGC project URL: http://mgc.nci.nih.gov
COMMENT      Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amgdbcm.tmc.edu
Guanarata, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.
FEATURES
Source
1..2067
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:30305 IMAGE:5134851"
/tissue_type="Liver, normal, 5 month old male mouse."
/clone_lib="NCI CGAP_L19"
/lab_host="DH10B"
/note="Vector: pCMV-Sport6"
10..1971
/codon_start=1
/product="unknown (protein for MGC:30305)"
/protein_id="AAH19376.1"
/db_xref="GI:18044173"
/translation="MGROAMWISSICPLPRPCFILLILLVLPVGAPOAGRNHTPEP
GNWMAVPMPTIPIVSGNSYTSRSPAEAFEGSPRYPDPSSSPGCOVLESG
OPRPFPRYGRMLHSTSGSNTRKWCATTHNDRBRAMKYCAEYTLVEGPAILD
CASPGLNGVCSSTHDHGSYHSCPLATGDCDETCDETRTYEFVEYGHMARS
EGHVDDCGMEGQARCEDHTHTACISSPLNGTCHLIVGTSVCTPGLTAGRRCN
IYVTEHFLMGTEYRGVASTAASGLSCLANMSDLVDELHVDVSAAYVLGIGRAY
GRHKRFLPRILIGSSSLPSGHPMLAATYIGNSPCAGSLVHTCWVYSAHCFANS
PRDSITVVLGHFPMFTDVTOTREIEXYVPTLYSVENPNHMDLYILRIKKGERC
AVRSQFVPLTPEAGSSPFTGKCOLIAGWGMENVSYSNLSLELYPLVADHKCS
SPRYGADISPMILCACTFECKSDACQDSSGSPLYCEKNYATLYGIIISMGDCGRIN
KGVITRYANTVMDINRIRPRPVATIS"
BASE COUNT      433 a 633 c 583 g 418 t
ORIGIN
Query Match      7.6%; Score 128.2; DB 10; Length 2067;
Best Local Similarity 50.8%; Pred. No. 5,5e-27;
Matches 620; Conservative 0; Mismatches 548; Indels 53; Gaps 11;

```

```

      469 AACCCCTGCCAGATGGGCTAAGTCTGCTCCCGCATAGGGAGATCCAACTTCACTGT 528
      742 AGCCCATGTCTGAACGAGGACACCTGATGTGGGCCACAGGGACACGCTGTCG 801
      529 GCGTGTCCACACAGTTCAGGGGAAATTTGTAAATAGT---TGTGATGACGCTAT 585
      802 ACCGAGCCGTTGGGCTATGCTGGGGTGTGTAACATTTGTTCCACAGACATCGCTTC 861
      586 GTTGGCATGGCTACTTACCAGAGGAAATGATAGACAGTCAACACAGATGCGTCG 645
      862 CTGGGAATGTGACAGATACGAGGCTGGCCAGCACCGCTGCGGCGCTGAGCTGC 921
      646 CTTACTGGAACCTCCACCTCTCTTGAGAGAAATTAACAACATGTTATGAGAGATCT 705

```

```

Db 922 CTGGCTGGAAATTCGTACCTGCTCTACCGAGACGTGACTCACTGCTGCTGCT 981
OY 706 GAAACCCATGGGATTTGGGAAACAAATTTCTGCAAGAAACCCAGATGGGACAAAGCCC 765
Db 982 GTTCGCTGTGGCTGGGCCCCCTACAGCTTACTGCTCCGGAACCGACAGAGATAGAGGCT 1041
OY 766 TTGGCTTTTAAAGTTACCAATGACAAAGTGAATGGAAATCTGTATGTCACCC 825
Db 1042 TGGTCTCTATGT---GGTGAAGGACAAACGACCTGTGGAGATTTGCGCTGACACCC 1098
OY 826 TGGTCAAGCCAGGAGCTTGGCTACCCAGAGAAAGCCCACTAGATCCATCAACCAAGCT 885
Db 1099 TGTGAATCCCTGGCCAGAGATCCACTCCCA-----AAGCCCGAGATCTTACGACCCCTG 1152
OY 886 CCGGGGTTTGAATCTGCTGGAAGAGTGAATGATGAGAGAGAGATCAAGAAATCTA- 944
Db 1153 CCGAGATCAAGCCCGGCTGTGCTCTTACTGTGTGACAGAGGACAAAGAGAGAGCTTC 1212
OY 945 TGGAGGCTTTAAGAGACGAGGCGGGAAGCAGCAGCAGGAGGCTGCTGCTGCTGCT 1004
Db 1213 TTGAGAGCAGCAGCATATCGGGGGGCTCTATCTTGTGCTGCTGCTGCTGCTGCTGCT 1272
OY 1005 GCGCTGACCATCTCATATGCCCCAGGGGCACTTGTGTGGGGGCTGATCCACCCCTG 1064
Db 1273 GCGATCTCATTTGGGAATA-----GCTTGTGTGCGGGAGGCTTGTCTATACCTG 1322
OY 1065 CTGGGTGCTCACTGCTGCCCATCTGACCGACATTAACAC-----AGACATCTAAAGCT 1118
Db 1323 CTGGGTAGTGTCTACGCCCATCTGCTTGTGCTCAACGACCCCGGAGGACGATCAAGT 1382
OY 1119 GGTGTAGGGGAGCAGGAGCTGGAAGAAAGAAATTTATGATAGAGAGCTTTAGGCTGA 1178
Db 1383 GGTAGTGGGTAGAGCATCTTCAACGCGCACAGGATGTGACAGAGCATTTGGCATTTGA 1442
OY 1179 GAAAGATTTCAAGTACAGCAGCATCAATGAAGAGATGATGCCCAAGATATGTC 1238
Db 1443 GAAAGATGTGGCTTACACCTGTACTGCTGTTCACACCCCA---ACACCATATGACTTGT 1499
OY 1239 ATTGCTCAAGTTAAAGCCAGTGGATGTCATGCTCTAGATATCCAAATAGCTGAAGAC 1298
Db 1500 CTGTGATCGGCTGAGAAAGAAAGAGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1559
OY 1299 TGTGTGCTTGGCTGATG-----GCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1352
Db 1560 CATGTGCTGCTGCTGAGGAGGAGGAGCTCTTCCCTGACACAAAGTGTGAGATTGCAAG 1619
OY 1353 CTGGGG-----TGTACAGAAACAGAAAGGCTCCCGCAGCTCTCTGATGATGCCAACT 1406
Db 1620 CTGGGAGCAGATGATGAAATGTGACAGCTACTTCAACTCTGCTGAGAGGAGCTGCT 1679
OY 1407 CAAGCTGATGGCAACACTTGTGCAACTCCCGCCAGCTCATATGACCATGATTTGATGA 1466
Db 1680 CCTCTTGTGTGCTGACCAAGAGTAGCAGGCCAGAGGATATATGCTGATCATACGCC 1739
OY 1467 CAGTATGATCTGTGAGAAATCTTGAAGAACTGGGCAACACCTGCGAGGAGCTAC 1526
Db 1740 TATCATGCTGTGTGCGGCTACTT---CGACTGCAAGTCCGATCCCTGCGGAGGAGACT 1796
OY 1527 TGGAGGCCCTTGAAGCTGTGAGAGAGAGGACACTACTACTGTATGATGATGAGCTG 1586
Db 1797 AGTGTGGGCTTGTGTGTGAGAAAGATGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1856
OY 1587 GGGCTGTGAGTGTGAG-----AAGAGGCGCAGGGGTCTACACCCAAATTTCAATTCCT 1640
Db 1857 GGGTATGAGCTGTGGGGGCTCAACAGCAGAGATCTACACCCGTGTGCTCAATTTATGT 1916
OY 1641 GAATTTGATCAAGACCACT 1661
Db 1917 GGACTGTATCAAGACGCTAT 1937

```

```

LOCUS AC106236 68586 bp DNA linear HTG 14-JUL-2002
DEFINITION Rattus norvegicus clone CH230-13J10, *** SEQUENCING IN PROGRESS
ACCESSION AC106236
VERSION AC106236.2 GI:21731435
KEYWORDS HTG, HTGS_PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 68586)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayala,M., Banks,T.,
Barbarta,J., Benton,J., Blimace,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Butay,C., Butch,P., Burrell,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Checko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Dayla,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Eamhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escoto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
Homs,J.F., Howard,S., Huber,J., Hulik,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlssoen,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,C.J., Kureishi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Louissege,H.,
Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,A., Ma,J.,
Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabadi,K., Morgan,M., Morris,S.,
Mosier,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okwundu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Plimus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G.,
Scherrer,S., Scott,G., Shen,H., Shooshari,N., Sisson,I.,
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Swalek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,D.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Welnstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 68586)
Direct Submission
Worley,K.C.
Submitted (12-JAN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 68586)
Worley,K.C.
Direct Submission
Submitted (14-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 11, 2002 this sequence version replaced gi:18138757.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information

```

```

Center project name: GKCO
Center clone name: CH230-13J10
----- Summary Statistics -----
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 29977 bases at least Q40
Consensus quality: 32280 bases at least Q30
Consensus quality: 34098 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.bgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
* NOTE: This is a "working draft" sequence. It currently
* consists of 41 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1160: contig of 1160 bp in length
1161
1280: gap of unknown length
1281
2765: contig of 1305 bp in length
2766
2865: gap of unknown length
2866
4036: contig of 1171 bp in length
4037
4136: gap of unknown length
4137
5378: contig of 1242 bp in length
5379
5478: gap of unknown length
5479
7017: contig of 1539 bp in length
7018
7117: gap of unknown length
7118
8400: contig of 1283 bp in length
8401
8500: gap of unknown length
8501
9562: contig of 1062 bp in length
9563
9662: gap of unknown length
9663
10706: contig of 1044 bp in length
10707
10806: gap of unknown length
10807
12186: contig of 1380 bp in length
12187
12286: gap of unknown length
12287
13687: contig of 1401 bp in length
13688
13787: gap of unknown length
13788
15164: contig of 1377 bp in length
15165
15264: gap of unknown length
15266
16319: contig of 1055 bp in length
16320
16419: gap of unknown length
16420
17529: contig of 1110 bp in length
17530
17629: gap of unknown length
17630
18947: contig of 1318 bp in length
18948
19047: gap of unknown length
19048
20129: contig of 1082 bp in length
20129
20229: gap of unknown length
20130
21241: contig of 1012 bp in length
21241
21341: gap of unknown length
21342
22617: contig of 1276 bp in length
22617
22717: gap of unknown length
22718
24183: contig of 1466 bp in length
24183
24283: gap of unknown length
24284
25879: contig of 1586 bp in length
25879
25979: gap of unknown length
25980
28141: contig of 2162 bp in length
28141
28241: gap of unknown length
28242
29800: contig of 1559 bp in length
29800
31624: gap of unknown length
31624
31724: contig of 1724 bp in length
31724
32926: gap of unknown length
32926
33026: contig of 1202 bp in length
33026
33227: gap of unknown length
33227
34376: contig of 1350 bp in length
34376
34476: gap of unknown length
34477
34777: gap of unknown length
34777
36187: contig of 1711 bp in length
36187
36287: gap of unknown length
36287
37398: contig of 1111 bp in length
37398
37498: gap of unknown length
37498
37599: gap of unknown length
37599
37499
39415: contig of 1917 bp in length
39415

```

```

* 39416 39515: gap of unknown length
* 39516 41278: contig of 1763 bp in length
* 41279 41378: gap of unknown length
* 41379 43355: contig of 1977 bp in length
* 43356 43455: gap of unknown length
* 43456 45020: contig of 1565 bp in length
* 45021 45120: gap of unknown length
* 45121 47073: contig of 1953 bp in length
* 47074 47172: gap of unknown length
* 47173 48525: contig of 1352 bp in length
* 48526 48625: gap of unknown length
* 48626 50339: contig of 1714 bp in length
* 50340 50439: gap of unknown length
* 50440 52494: contig of 2055 bp in length
* 52495 52594: gap of unknown length
* 52595 54163: contig of 1569 bp in length
* 54164 54263: gap of unknown length
* 54264 56386: contig of 2123 bp in length
* 56387 56486: gap of unknown length
* 56487 58244: contig of 1758 bp in length
* 58245 58344: gap of unknown length
* 58345 60191: contig of 1847 bp in length
* 60192 60291: gap of unknown length
* 60292 62520: contig of 2229 bp in length
* 62521 62621: gap of unknown length
* 62621 65096: contig of 2476 bp in length
* 65097 65196: gap of unknown length
* 65197 68586: contig of 3390 bp in length.

```

```

FEATURES
    source          1..68586
                    /organism="Rattus norvegicus"
                    /db_xref="taxon:10116"
                    /clone="CH230-13J10"

```

```

BASE COUNT      15651 a 16674 c 16111 g 16135 t 4015 others
ORIGIN

```

```

Query Match          7.1%; Score 119; DB 2; Length 68586;
Best Local Similarity 82.0%; Pred. No. 5.3e-24;
Matches 137; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

```

```

OY 1517 AGGCTGACTGTGGAGGCCCCCGACCTGTGAGGAGCGGACGACTACTAGCTTATGGA 1576
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 53672 AGGCTGACTGTGGAGGCCCCCTTAACTGTGAGGAGGTGAACTTACTACTGTATGGA 53613
OY 1577 TAGTGAGCTGGGCTCGAGTGTGAGGAGGCGGCTTACACCCCAAGTTTCAAAAT 1636
    ||||||| ||||||| || |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB 53612 TTCTAAAGCTGGGGCCAGGAACTGTGGAGGAGGCTGAGCTCAAGTCAAGTCAAGT 53553
OY 1637 TCCGTAATGTGATCAAGCCACCATCAAAAGTGAAGTGGCTTCTTAA 1683
    ||||||| ||||||| || |||| |||| || || || |||| |||| |||| |||| ||||
DB 53552 TCCTGAATTGGATMAAACCAACCATGCAACAGGAGGCTGCTCTGA 53506

```

```

Search completed: March 6, 2003, 18:44:14
Job time : 4882 secs

```

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 15:20:13 ; Search time 2178.5 Seconds

(without alignments)
12511.822 Million cell updates/sec

Title: US-09-912-559-2

Perfect score: 1683

Sequence: 1 aggttcgcagatgtctga.....aaagtgaagtgtctctaa 1683

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154056 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estro:*
9: gb_estl:*
10: gb_estl2:*
11: gb_estl3:*
12: gb_estl4:*
13: gb_estl5:*
14: gb_estfun:*
15: em_estfun:*
16: em_estcom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_frod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	698.2	41.5	891	13	BT761782 603046775
2	451.6	26.6	892	10	BE032018 BE032018
3	447.2	25.9	802	13	BC972579 602841165
4	436.2	25.9	674	9	AA217892 mvs4b08.r
5	435.4	25.9	751	13	BT148082 602912432
6	432	25.7	570	13	BM508620 1137h03.y

7	416.4	24.7	929	12	BT788188
8	396.4	23.6	791	13	BT332440
9	394.2	23.4	546	9	AA237499
10	392.6	23.3	608	10	AV601564
11	390.4	23.2	902	12	BT780971
12	386	22.9	682	13	BT220028
13	384.6	22.9	929	12	BT785781
14	372.4	22.1	969	12	BT384535
15	370.6	22.0	654	10	AM475402
16	370.4	22.0	587	13	BG972681
17	365.6	21.7	659	10	AM610902
18	363.6	21.6	594	13	BM503097
19	347.2	20.6	716	12	BT789705
20	335.8	20.0	655	10	BM569555
21	330.8	19.7	934	13	BT765113
22	327.8	19.5	345	14	T68666
23	317.4	18.9	486	9	AA286125
24	311.8	18.5	600	12	BT788541
25	302.4	18.0	335	14	R89458
26	293.2	17.4	304	14	T77362
27	279.8	16.6	521	13	BT221626
28	275.8	16.4	912	13	BG972626
29	274	16.3	366	14	T93666
30	264.4	15.7	434	14	T84369
31	258.6	15.4	781	12	BT533788
32	256.2	15.2	458	14	R10295
33	245.4	14.6	719	13	BT102281
34	243.8	14.4	593	9	AT035669
35	241.8	14.4	513	13	BT247736
36	241	14.3	384	14	W66981
37	241	14.3	397	9	AA002504
38	239.8	14.2	353	14	R11014
39	235.6	14.0	489	9	AT787532
40	235.4	14.0	393	14	T50880
41	213.4	12.7	530	10	BE235171
42	204.8	12.2	361	9	AA108230
43	197.8	11.8	320	14	T69534
44	151.8	9.0	642	13	BT067550
45	150.4	8.9	810	9	AT255992

ALIGNMENTS

RESULT 1
BT761782
LOCUS
DEFINITION
603046775f1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:518706 5',
mRNA sequence.
BT761782
ACCESSION
BT761782.1 GI:15753360
VERSION
BT761782
KEYWORDS
EST.
SOURCE
ORGANISM
Homo sapiens
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 891)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LHAM1467 row: e column: 19
High quality sequence stop: 839.
Location/Qualifiers
1. 891

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5187066"
/clone_lib="NIH_MGC_116"
/lab_host="DH10B"
/Note="Organ: pooled colon, kidney, stomach; Vector:
pcmw-sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-qt primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH_MGC Library."

BASE COUNT      243 a      234 c      211 g      203 t
ORIGIN

Query Match      41.5%; Score 698.2; DB 13; Length 891;
Best Local Similarity 97.2%; Pred. No. 7.6e-188;
Matches 775; Conservative 0; Mismatches 13; Indels 9; Gaps 6;

OY 1 ATGTTGCCAGATGCTGTGATCTCCATGTTGCTGTTAATGGCTCTGGTGGAAACA 60
    |||||||
DB 81 ATGTTGCCAGATGCTGTGATCTCCATGTTGCTGTTAATGGCTCTGGTGGAAACA 140
    |||||||
OY 61 GCGGTGGGTTCCCTGATGCTTATTTGGAAGCCTGGACCCAGATGACCCCTGAC 120
    |||||||
DB 141 GCGGTGGGTTCCCTGATGCTTATTTGGAAGCCTGGACCCAGATGACCCCTGAC 200
    |||||||
OY 121 CAGTATGATTACAGTACGAGGATTAATATCAGAGAAACACGATAGCACCTTACC 180
    |||||||
DB 201 CAGTATGATTACAGTACGAGGATTAATATCAGAGAAACACGATAGCACCTTACC 260
    |||||||
OY 181 CATGCTGAGATCCTGATGCTACTACTACAGAGACCATGATGATCCAGCCCAAC 240
    |||||||
DB 261 CAGCTGAGATCCTGATGCTACTACTACAGAGACCATGATGATCCAGCCCAAC 320
    |||||||
OY 241 CCTGTGAAACGCTGGGAGACTGCTCTCCATGAGGACACTTCAATCAGCTGCTG 300
    |||||||
DB 321 CCTGTGAAACGCTGGGAGACTGCTCTCCATGAGGACACTTCAATCAGCTGCTG 380
    |||||||
OY 301 GCTCCTTCTCTGGGAATAAGTGTCAAGAAATGCAAAATACGTGCAAGAGCAACCATGT 360
    |||||||
DB 381 GCTCCTTCTCTGGGAATAAGTGTCAAGAAATGCAAAATACGTGCAAGAGCAACCATGT 440
    |||||||
OY 361 GCGCGGGGCCAATGCTCATTTACCCAGAGTCTCTCCCTTACTACCGCTGTGTTAAAC 420
    |||||||
DB 441 GCGCGGGGCCAATGCTCATTTACCCAGAGTCTCTCCCTTACTACCGCTGTGTTAAAC 500
    |||||||
OY 421 CCTTACAGAGTCCAGACTGCTCCCAAGTGTGTTCTGTATGACAGGCCAACCCTGCGAC 480
    |||||||
DB 501 CCTTACAGAGTCCAGACTGCTCCCAAGTGTGTTCTGTATGACAGGCCAACCCTGCGAC 560
    |||||||
OY 481 AATGGGGCTACTCTCTCCCGGCAATAGCGAGATCAAGATTCACTGTGCTGCCGAC 540
    |||||||
DB 561 AATGGGGCTACTCTCTCCCGGCAATAGCGAGATCAAGATTCACTGTGCTGCCGAC 620
    |||||||
OY 541 CAGTTCAAGGGGAATCTGTGAATAGTTCGATGACTGCTATGTTGGCGATGGCTAC 600
    |||||||
DB 621 CAGTTCAAGGGGAATCTGTGAATAGTTCGATGACTGCTATGTTGGCGATGGCTAC 680
    |||||||
OY 601 TCTTTACCAAGGAAATGATAGACAGTTC - AACGAGATGCGTCTTACTGSAATC 659
    |||||||
DB 681 TCTTTACCAAGGAAATGATAGACAGTTC - AACGAGATGCGTCTTACTGSAATC 740
    |||||||
OY 660 CC - ACCCTCTTTCAGAGAAATTAACACATGTTATGAGG - -ATGCGAAACCATG 715
    |||||||
DB 741 CCAAGCTCTCTTCAGAGAAATTAACACATGTTATGAGGAGATGCTTAAACCATG 800
    |||||||
OY 716 GGAATGGGAA - CACAATTTTCGAG - AAACCCAGATGGGAGG - AAAGCCCTGCTGC 771
    |||||||

```

```

DB 801 GGATTGGGAACCAATTTCTGCAAGAAACCCAGATGCCGACGAAAAAGCCCTGGTGC 860
OY 772 TTATTTAAAGTACCAA 788
    |||||||
DB 861 TTATTTAAAGTACCA 877

RESULT 2
BE032018
LOCUS BE032018 592 bp mRNA linear EST 09-JUL-2000
DEFINITION 130809 MARC IPiG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BE032018
VERSION BE032018.1 GI:8327027
KEYWORDS EST.
SOURCE Sus scrofa
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 592)
Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grose,W.M., Bennett,G.A., Laegreid,W.W.
and Keele,J.W.
Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
Unpublished (2000)
JOURNAL Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68935-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt-trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCCAGTACGACG
Plate: 69 row: K column: 12
Seq primer: ATTTAGGTGACACTAATG.
FEATURES
source
location/Qualifiers
1..592
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC IPiG"
/tissue_type="pooled"
/lab_host="DH10B"
/Note="Vector: PCMV SPORT6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

BASE COUNT      144 a      152 c      174 g      122 t
ORIGIN

Query Match      26.8%; Score 451.6; DB 10; Length 592;
Best Local Similarity 86.0%; Pred. No. 1.5e-117;
Matches 518; Conservative 0; Mismatches 69; Indels 15; Gaps 1;

OY 522 CACCTGTGCTGTGCCGACCGATTCAAGGGGAATTCGTGAAATAGTTCGTGATGAC 581
    |||||||
DB 6 CCGGTGGGCTGTGCTGTGACCAAGTTAAAGGAGATTCGTGAAATAGTTCGTGATGAC 65
    |||||||
OY 582 CTATGTGGCATGCTCTCTTACCCAGGAAATGATAGACAGTCAACGACGATGC 641
    |||||||
DB 66 CTATGTGTGGAGCGGCTACTCTTACGAGGGGAAAGTAAACTGTCAACGACGAC 125
    |||||||
OY 642 GTGCTTTTACTGGAATCCCACTCTCTTTCAGAGGAATTAACAATGTTTATGAGGA 701
    |||||||
DB 126 GTGCTTTTACTGGAATCCCACTCTCTTTCAGAGGAATTAACAATGTTTATGAGGA 185
    |||||||
OY 702 TCGTGAACCATGGGATTTGGGAGACCAATTTCTGACAGAACCCAGATGGGAGAAA 761
    |||||||
DB 186 TGCAGAGCCCATGGATTTGGGAGACCACTTCTGCAGAAACCCAGATGGAGACAGAAA 245
    |||||||
OY 762 GCCTGTGCTTATTAAGTATACCAATGACAAAGTGAATGGGAATACTGTGATGCTC 821
    |||||||

```



```

OY 123 GTATGATACGATGAGAGATTATTAATCAGAAAGAACACCAAGTACGACACTTACCA 182
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 122 CTATATACAGCTATGACAGATCCAGCCAGAGAGAACCCAGTGTCCAGCAGACAC 181
OY 183 TGTGAGAAATCCTGACTGTACTACACTGAGACCAAGCTGATCCATGCCAGCCAAACC 242
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 182 CCGTAGAGAACCCCGAGTGTACT-----ATGAAAGAGATGATCATGCCAGTCCACCC 235
OY 243 CTGTGAACAGAGGGGAGAGCTGCTGTGCATGGAGAGACCTTCACATGAGCTGCTGGGC 302
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 236 CTGTGAACAGAGGGGAGAGCTGCTGTGCATGGAGAGACCTTCACATGAGCTGCTGGGC 295
OY 303 TCCCTTCTGTGGGAATTAAGTGTGCAAGAAATGCAAAATACGTGCAGAGCAACCCATGTGG 362
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 296 CCCCTTCTGTGGGAGCCGCTGTGCAGAGCTGCACAAACAAAGTGTGAGAGCAACCCATGTGT 355
OY 363 CCGGAGGCATATGCTCATACAGAGCTCCCTACTACCTGCTGTGTGTAACACCC 422
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 356 CCATGCTGATGCTGCTATTTACCCAGAAACACCCCTACTACCTGCTGTGTGTAACATGCC 415
OY 423 TTACACAGTCCAGCTGCTCCCAAGTGTCTGTATGACAGGCAAAACCCCTGCCAGAA 482
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 416 TTACACAGGACCCAGACTGTCTCCAAAGTGTCTCCGAGTGTGAGCCCAAAACCCCTGCCAGAA 475
OY 483 TGGGGCTACCTGCTGCCGAGATACCGGAGATCCCAAGTTCACCTGTGCTGCCAGCA 542
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 476 TGGGGAGTGTGTTCCCGACACAGAGAGATCCAGTGTACTGTGCTGTGCCAGACCA 535
OY 543 GTTCAAGGGAATTTCTGGAATATAGTTCTGATGACTGCTATGTTGGGATGCTACTC 602
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 536 GTATTAAGGGAATTTCTGGAATATAGTTCTGATGACTGCTATGTTGGGATGCTACTC 595
OY 603 TTACAGGAGAAATTAATAGACAGTCAACAGCATGCTGCTTACTTGG-NACTGCC 661
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 596 TTACAGGAGAAATTAATAGACAGTCAACAGCATGCTGCTTACTTGG-NACTGCC 655
OY 662 ACCTGCTTGTGACGAGAAATTAACATGTTTATGAGAGTGTGAACCCATGSGAATTG 721
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 656 ACCCTCTTGTGACGAGAAATTAACATGTTTATGAGAGTGTGAACCCATGSGAATTG-TC 714
OY 722 GGAAGACAAATTTCTGCAAGAACCCAGATGGCGAG 756
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 715 GAGAGACAACTTCTG-AGAAACCCAGATGGAGAC 748
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
BM508620 570 bp mRNA linear EST 12-MAR-2002
LOCUS 1137H03.y1 Melton Normalized Mixed Mouse Pancreas 1 N1-MMS1 Mus
DEFINITION musculus cDNA clone IMAGE:5944156 5' similar to TR:Q14520 Q14520
GCF ACTIVATOR LIKE PROTEIN. ; mRNA sequence.
ACCESSION BM508620
VERSION BM508620.1 GI:18679763
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 570)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishke,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
,M., Gibbons,M., McCann,R., Cole,R., Tsagaris,I., Williams,T.,
Jackson,Y., and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Other ESTs: 1137H03.x1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138

```

```

FEATURES
    source
        1..570
            /organism="Mus musculus"
            /strain="ICR"
            /db_xref="taxon:10090"
            /clone="IMAGE:5944156"
            /clone_1fb="Melton Normalized Mixed Mouse Pancreas 1
            N1-MMS1"
            /sex="Both for embryonic & newborn, male for adult and
            adult Islet."
            /dev_stage="Embryonic day 10.5, E12.5, E16.5, newborn,
            adult, mixed"
            /lab_host="DH10B"
            /note="Vector: pSPORT1; Site.1: Not I; Site.2: Sal I; Five
            libraries representing E10.5/12.5 pancreatic bud, E16.5
            pancreas, newborn pancreas, adult pancreas, and adult
            islets of Langerhans were separately constructed using
            Superscript Plasmid Library Kit (Life Technologies). cDNA
            was made by oligo-dT priming and size-selected by column
            fractionation. Libraries were amplified once on solid
            support and plasmid DNA from each library was prepared
            and mixed in equal amounts. The mixed library DNA was
            normalized by method #4 from Bonaldo, Lennon, and Soares
            1996 Genome Research 6:791-806. 0.5 microgram
            single-stranded mixed library plasmid DNA was mixed with
            5 micrograms PCR product representing mixed library
            inserts and hybridized to an EcoT of 6. Single-stranded
            (unhybridized) plasmids were isolated by hydroxyapatite
            chromatography and used to make this library."
BASE COUNT 162 a 134 c 152 g 122 t
ORIGIN
Query Match 25 7%: Score 437: DB 113: Length 570:
Best Local Similarity 85.0%: Pred. No. 5.8e-112:
Matches 483: Conservative 0: Mismatches 85: Indels 0: Gaps 0:
OY 1108 CATCTAAGGTGCTGCTAGGGAGCAGACCTGTAAGAAAGAAATTTTCATGACAGAC 1167
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3 CATCTAAGGTGCTGCTAGGGAGCAGACCTGTAAGAAAGAAATTTTCATGACAGAC 62
OY 1168 TTTAGGCTGCAGAAATTAATCAATGAGCAATGAAAGAGATGGATTTCCCGAC 1227
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 63 TTTAGGCTGCAGAAATTAATCAATGAGCAATGAAAGAGATGGATTTCCCGAC 122
OY 1228 AATGATATTGCAATTTGCTCAAGTTAAAGCAGTGGATGTGCTGCTTGAATCCAA 1287
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 123 AATGATATTGCTTCTCAATGTTAAAGCAGTGGATGTGCTGCTTGAATCCAA 182
OY 1288 TAGCTGAACAGTGTGCTGCTGATGAGGCTCTTCCCTGGGAGGTGGCCATC 1347
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 183 TAGCTGAACAGTGTGCTGCTGATGAGGCTCTTCCCTGGGAGGTGGCCATC 242
OY 1348 TCTGGCTGGGCTTTACAGAAACAGAAAGGCTCCCGCAGCTCTGATGCCAAATC 1407
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 243 TCTGGCTGGGCTTTACAGAAACAGAAAGGCTCCCGCAGCTCTGATGCCAAATC 302
OY 1408 AAGCTATTGGCAACACTTTGTGCAACTCCCGCAGCTCTGATGCCAAATGATGC 1467
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 303 AAGCTATTGGCTTAACCTTTGTGCAACTCCCGCAGCTCTGATGCCAAATGATGC 362
OY 1468 AGTATGATCTGTGACGAAATCTTACAGAAACCTGGGCAAGACACCTGCGAGGCTGT 1527
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Tel: 617-495-1812
Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:

Washington University Genome Sequencing Center for information on
obtaining a clone please contact: Juliana Brown
(brown@fas.harvard.edu)

GI:2007028 This sequence now available from the IMAGE consortium,
for clone orders contact: info@image.llnl.gov

Seq primer: -40RP from Gibco

High quality sequence stop: 435.

FEATURES

source

1..570

/organism="Mus musculus"

/strain="ICR"

/db_xref="taxon:10090"

/clone="IMAGE:5944156"

/clone_1fb="Melton Normalized Mixed Mouse Pancreas 1

N1-MMS1"

/sex="Both for embryonic & newborn, male for adult and

adult Islet."

/dev_stage="Embryonic day 10.5, E12.5, E16.5, newborn,

adult, mixed"

/lab_host="DH10B"

/note="Vector: pSPORT1; Site.1: Not I; Site.2: Sal I; Five

libraries representing E10.5/12.5 pancreatic bud, E16.5

pancreas, newborn pancreas, adult pancreas, and adult

islets of Langerhans were separately constructed using

Superscript Plasmid Library Kit (Life Technologies). cDNA

was made by oligo-dT priming and size-selected by column

fractionation. Libraries were amplified once on solid

support and plasmid DNA from each library was prepared

and mixed in equal amounts. The mixed library DNA was

normalized by method #4 from Bonaldo, Lennon, and Soares

1996 Genome Research 6:791-806. 0.5 microgram

single-stranded mixed library plasmid DNA was mixed with

5 micrograms PCR product representing mixed library

inserts and hybridized to an EcoT of 6. Single-stranded

(unhybridized) plasmids were isolated by hydroxyapatite

chromatography and used to make this library."

BASE COUNT

162 a

134 c

152 g

122 t

ORIGIN

Query Match

25 7%: Score 437: DB 113: Length 570:

Best Local Similarity 85.0%: Pred. No. 5.8e-112:

Matches 483: Conservative 0: Mismatches 85: Indels 0: Gaps 0:

OY 1108 CATCTAAGGTGCTGCTAGGGAGCAGACCTGTAAGAAAGAAATTTTCATGACAGAC 1167

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

DB 3 CATCTAAGGTGCTGCTAGGGAGCAGACCTGTAAGAAAGAAATTTTCATGACAGAC 62

OY 1168 TTTAGGCTGCAGAAATTAATCAATGAGCAATGAAAGAGATGGATTTCCCGAC 1227

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

DB 63 TTTAGGCTGCAGAAATTAATCAATGAGCAATGAAAGAGATGGATTTCCCGAC 122

OY 1228 AATGATATTGCAATTTGCTCAAGTTAAAGCAGTGGATGTGCTGCTTGAATCCAA 1287

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

DB 123 AATGATATTGCTTCTCAATGTTAAAGCAGTGGATGTGCTGCTTGAATCCAA 182

OY 1288 TAGCTGAACAGTGTGCTGCTGATGAGGCTCTTCCCTGGGAGGTGGCCATC 1347

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

DB 183 TAGCTGAACAGTGTGCTGCTGATGAGGCTCTTCCCTGGGAGGTGGCCATC 242

OY 1348 TCTGGCTGGGCTTTACAGAAACAGAAAGGCTCCCGCAGCTCTGATGCCAAATC 1407

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

DB 243 TCTGGCTGGGCTTTACAGAAACAGAAAGGCTCCCGCAGCTCTGATGCCAAATC 302

OY 1408 AAGCTATTGGCAACACTTTGTGCAACTCCCGCAGCTCTGATGCCAAATGATGC 1467

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

DB 303 AAGCTATTGGCTTAACCTTTGTGCAACTCCCGCAGCTCTGATGCCAAATGATGC 362

OY 1468 AGTATGATCTGTGACGAAATCTTACAGAAACCTGGGCAAGACACCTGCGAGGCTGT 1527

||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 363 AGTATGATTGTGGGGGAACCTTCAGAAAGCCCGATCAGACACCTGCCAGGGTGAATCG 422
QY 1528 GGAGGGCCCCCTGACCTGTGAGAGAGAGGACGACCTACTACTGCTATGGGATAGAGACTGG 1587
Db 423 GGGGGCCCTCTACCTGTGAGAGAGATGAGACTTACTGCTGCTGGGATTTGTAACCTGG 482
QY 1588 GGCTGTGAGTGTGAGAGAGGCGGCTCTACACCCAGTGTACCAATTTCTGAATTGG 1647
Db 483 GGCGAGGATGTGGAGAGAGGAGGAGTCTTACACTCAAGTACACCAAGTTCCTGATTTGG 542
QY 1648 ATCAAGGACCACTCAAGTAAAGTGTG 1675
Db 543 ATTAAGACCACTATGACAGGAGGAGCTG 570

RESULT 7
BF78188
LOCUS 60213411F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:4241642
DEFINITION 5' mRNA sequence.
ACCESSION BF78188 929 bp mRNA linear EST 12-JAN-2001
VERSION BF78188
KEYWORDS GI:12093224
EST.
SOURCE house mouse.
ORGANISM Mus musculus.

REFERENCE 1
AUTHORS NIH-MGC
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9858 row: e column: 03
High quality sequence stop: 686.

FEATURES

source
1..929
Location/Qualifiers
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4241642"
/clone_lib="NCI_CGAP_Kid14"
/lab_host="DH10B (TI phage-resistant)"
/note="Organ: Kidney; Vector: pCMV-Sport6; Site: 1; NotI; Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library. 1"
BASE COUNT 238 a 275 c 253 g 162 t 1 others
ORIGIN

Query Match 24.7%; Score 416.4; DB 12; Length 929;
Best Local Similarity 75.9%; Pred. No. 2,1e-107;
Matches 529; Conservative 0; Mismatches 162; Indels 6; Gaps 1;

QY 1 ATGTTGGCAGGATGCTGATCCATGCTGTTAATGCTGTGGTGGAAAGACA 60
Db 44 ATATTGTGACAGGATGAGTGTTCGCTGCTGCTAATCCCTGGTGGGAAGTCA 103
QY 61 GCGTGTGGTTCCTCCGATGCTTATTGGAAGCCTGGAGCCAGACAGCCCTGAC 120
Db 104 GTCATGGGGGCTTCACGATGCTTCATTTGCGCCCAACACCAATGAGACCCCAT 163
QY 121 CAGTATGATTACGCTACGAGATTAATATAGAGAGAAACCACTAGCACTTACC 180
Db 164 GACTATTACTACGCTATGAGCATGCCAGCCAGACGAAGACCCAGTGTCTACGAGACC 223

QY 181 CATGCTGAGATCTGACTGCTACTACACTGAGAGACCAAGCTGATCCATGCCACCCAC 240
Db 224 ACCCTGAGAACCCCGACAGTGAATGACT-----ATGAAGAGATGATTCATGCACTGAC 277
QY 241 CCCTGTGAACAGGAGGAGGAGCTGCTGCTCATGAGGAGACCTTCATGATGAGTGGCTG 300
Db 278 CCCTGTGAACAGGAGGAGGAGCTGCTGCTCATGAGGAGGAGTACCTTGAATGAGTGGCCA 337
QY 301 GCTCTTCTCTGAGAAATAGTGTGACAAAGTGTCAAAATAGTGTCAAGAGACACCTGTT 360
Db 338 GCGCCCTCTCGGGAGCCGGGTGCGAGACTGCACAAACAAAGTGTGAGAGAACCCATGT 397
QY 361 GGCGGGGCGCAATGCTCATATTCACAGAGCTGCTTACCTGCTGCTGTGTAAACAC 420
Db 398 GTCCATGATGATTTGCTCATTTACCCAGAGACACCTTACTACGCTGTGCTGTCAATAC 457
QY 421 CTTTACACAGTCCAGCTGCTCCCAAGTGTCTGTGTGAGGCAAAACCTGCGCAG 480
Db 458 CTTTACACAGGAGACAGCTGCTCCAAAGTGTCTCGGATGTGACAGGCCAAACCTGCGCAG 517
QY 481 AATGGGCTACCTGCTCCCGCATTAACGAGATCCAAAGTCACTGCTGCTGTCCGAC 540
Db 518 AATGGGAGAGTGTGTTCCCGACACAGAGAGATCCAGGTTTACCTGTGCTGTCCAGAC 577
QY 541 CAGTTCAGGAGAAATTTCTGGAATAGTGTCTGATGACTGCTATGTTGGCGATGCTAC 600
Db 578 CAGATTAAGGAGAAATCTGTGAATAGTGTGCGGAGCACTGTATGTGGTGGATGCTAC 637
QY 601 TCTTACGAGGAGAAATGATAGAGACATCAACAGCACTGCTGCTTACTGAGACTCC 660
Db 638 TCTTACGAGGAGCAAGTGTGATAGACAGTCCACCAAGCATGCTTACTGAGACTCC 697
QY 661 CACCTCTCTGAGAGAAATTAACATGATTAATGG 697
Db 698 CACCTCTCTGAGGAGACTTATACATGATGCTCG 734

RESULT 8
BI332440 791 bp mRNA linear EST 30-JUL-2001
LOCUS BI332440
DEFINITION 602980841F1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:5133587 5',
mRNA sequence.
ACCESSION BI332440 791 bp mRNA linear EST.
VERSION BI332440.1 GI:15017097
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.

REFERENCE 1
AUTHORS NIH-MGC
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM11328 row: a column: 12
High quality sequence stop: 787.

FEATURES

source
1..791
Location/Qualifiers
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5133587"
/clone_lib="NCI_CGAP_L19"
/lab_host="DH10B (TI phage-resistant)"
/note="Organ: liver; Vector: pCMV-Sport6; Site: 1; NotI;


```

Db 424 TATGTGAAGACTATGTTTGGCCAGCACCCTTTTCCCTCTGGAACGTAGTGGCACATC 483
OY 1348 TCTGTGGGGGTGTACAGAAACAGGAGGTCCGCCAGCTCCTGTGATGCCAAAGTC 1407
Db 484 TCTGTGGGGGTGTACAGAAACAGGAGGTCCGCCAGCTCCTGTGATGCCAAAGTC 543
OY 1408 AAG 1410
Db 544 AAG 546

RESULT 10
AV601564 608 bp mRNA linear EST 27-NOV-2001
AV601564 Bos taurus kidney fetus Bos taurus cDNA clone E1K1005A04
LOCUS 5', mRNA sequence.
DEFINITION AV601564
ACCESSION AV601564.1 GI:9723884
VERSION AV601564.1
KEYWORDS EST.
SOURCE Bos taurus
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 608)
Takasuga,A., Hirotsune,S., Itoh,R., Jitohzono,A., Suzuki,H., Aso,H.
and Sugimoto,Y.
Establishment of a high throughput EST sequencing system using
poly(A) tail-removed cDNA libraries and determination of 36,000
bovine ESTs
Nucleic Acids Res. 29 (22), E108 (2001)
21570554
Contact: Yoshikazu Sugimoto
Animal Genetics Division
Shirakawa Institute of Animal Genetics
Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
Tel: 81-248-25-5641
Fax: 81-248-25-5725
Email: kazusugie@ccocn.ocn.ne.jp
Single pass sequencing.
This clone was obtained from a polyA-deleted cDNA library.
Location/Qualifiers
1..608
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone="E1K1005A04"
/clone_lib="Bos taurus kidney fetus"
/tissue_type="kidney"
/dev_stage="fetus"
/lab_host="DH10B"
/Note="Vector: pML1, Site_1: SalI; Site_2: NotI; Poly A
was deleted from a NotI site"
BASE COUNT 148 a 171 c 151 g 135 t 3 others
ORIGIN
Query Match 23.3%; Score 392.6; DB 10; Length 608;
Best Local Similarity 81.9%; Pred. No. 1,1e-100;
Matches 465; Conservative 0; Mismatches 97; Indels 6; Gaps 1;

OY 1 ATGTTTGCCAGAGTCTGATCTCATCTGCTGTTAATGGCTCTGTGGGAAAGACA 60
Db 47 AATGTTTGCCAGAGTCTGATCTCATCTGCTGTTAATGGCTCTGTGGGAAAGACA 106
OY 61 GCGGTGGGTTCGCCGATGCTTTAATGGAAAGCCGACCCAGAGTGGACCCCTGAC 120
Db 107 GCGTTCGGGCTCTCCCTGCTGCTTTTCTCACAGAGCCGAGACCCAAATTGGACTCTCTAC 166
OY 121 CAGTAGATTACAGACTAGAGATTAAATCAGAGAGAACACCACTAGCACTTACC 180
Db 167 CAGTAGATTACAGACTAGAGATTAAATCAGAGAGAACACCACTAGCACTTACC 226
OY 181 CAGTAGATTACAGACTAGAGATTAAATCAGAGAGAACACCACTAGCACTTACC 240

```

```

Db 227 TATTCGACACACCTGAGCTGATCTACAGAGAGA-----TGACCATCCCTGTCCAG 280
OY 241 CCCTGTGACACAGCTGGGAGACTGCTGTGCATCCATGGAGACCTTCACATGACGTGCTCG 300
Db 281 CCCTGTGACACAGCTGGGAGACTGCTGTGCATCCATGGAGACCTTCACATGACGTGCTCG 340
OY 301 GCTCCTCTCTGCGGAATTAAGTGCAGAAAGTGCAGAAATACGTGCAGAGACACCACTG 360
Db 341 GACCTTTCTCTGGAACACAGTGTGACAGATGTGCAAAACAGTGCAGAAACACCACTG 400
OY 361 GCGCGGGGCAATGCTTCATATTACCCAGAGTCTCCTTACTACCGCTGTGTCTAATAC 420
Db 401 GCGCGGGGCAATGCTTCATATTACCCAGAGTCTCCTTACTACCGCTGTGTCTAATAC 460
OY 421 CCTTACAGAGTCCAGAGTGTCCCAAGTGGTTCCTGATACAGAGCAACCCCTGAC 480
Db 461 CCTTACAGAGTGTCCAGAGTGTCCCAAGTGGTTCCTGATACAGAGCAACCCCTGAC 520
OY 481 AATGGGCTACTGCTGCCCGCATTAACGAGATCCAAAGTTCACCTGCTGCTCCGAC 540
Db 521 AATGGGCTACTGCTGCCCGCATTAACGAGATCCAAAGTTCACCTGCTGCTCCGAC 580
OY 541 CACTTCAGAGGAATTCGTGGAATAG 568
Db 581 CACTTCAGAGGAATTCGTGGAATAG 608

RESULT 11
BF780971 902 bp mRNA linear EST 12-JAN-2001
LOCUS 602105493P1 NC1_CGAP_Kid14 Mus musculus cDNA clone IMAGE:4223710
DEFINITION 5', mRNA sequence.
ACCESSION BF780971
VERSION BF780971.1 GI:12086004
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 902)
NIH-MGC http://mgc.ncl.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c94pbs@email.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM9812 row: 1 column: 23
High quality sequence stop: 717.
Location/Qualifiers
1..902
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4223710"
/clone_lib="NC1_CGAP_Kid14"
/lab_host="DH10B (71 phage-resistant)"
/Note="Organ: kidney; Vector: PCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Constructed by Life
Technologies. Note: this is a NC1_CGAP Library. I"
BASE COUNT 223 a 264 c 254 g 161 t
ORIGIN
Query Match 23.2%; Score 390.4; DB 12; Length 902;
Best Local Similarity 73.1%; Pred. No. 5,5e-100;
Matches 532; Conservative 0; Mismatches 186; Indels 10; Gaps 2;

```

[illegible]

	cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov Plate: LHAM11236 row: 1 column: 03 High quality sequence start: 2 High quality sequence stop: 682. Location/Qualifiers
FEATURES	1..682
source	/organism="Mus musculus" /strain="FVB/N" /db_xref="taxon:10090" /clone="IMAGE:5098442" /clone_lib="NCI_GCAP_L19" /lab_host="DH10B (TI phage-resistant)" /note="Organ: liver; Vector: pcwv-SPORT6; Site:1; Notcl: Site2; Salt: Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI_GCAP library."
BASE COUNT	166 a 206 c 169 g 141 t
ORIGIN	
Query Match	22.9%; Score 386; DB 13; Length 682;
Best Local Similarity	77.0%; Pred. No. 8.e+99;
Matches 511; Conservative 0; Mismatches 145; Indels 8; Gaps 3;	
OY 1	ATGTTGCCAGGATGTCTGATTCACATGTCCTGTTAATGGCTCTGTGGGAACA CA 60
Db 24	ATTATTGTGAGATGTGTGTTCCTCGTCCGCTGCTAATGCCCTGTGGGAGCA 83
OY 61	GCCGTGGTCTCCCTCATGTCTTTATTGGAAGCCTTGACCAGACCTGACCCCTGAC 120
Db 84	GTCATTGGGCTCTCACTGATGTCCTTCATCTGGCCCCAACACCAATATGGACCCCGAT 143
OY 121	CAGTATGATTACACTGACGAGGATTTATATACGAGAAGAACCGATGACCACTTA AC 180
Db 144	GACTATTACTACAGCTATGACAGCTCACGCCAGACAAACCCAGTTCACGACAG CC 203
OY 181	CATGCTGAGAACTCTGACTGTACTACACTGAGACCAAGCTGATCATGCCACCA C 240
Db 204	ACCCCTGAAACCCCGCACTGTGTACT-----ATGAACACGATGATCATGCCAGTCCA AC 257
OY 241	CCCTGTGAACACGGGTGGGAGCTCCGTCGTCATGGGAGACCTTCACATGACGCTG CTG 300
Db 258	CCCTGTGAACACGGCGGGGACTGTATCATACAGAGGGATCTTCAAGTTGACGCTGCC A 317
OY 301	GCTCTCTTCTCTGGCAATTAAGTGTCAAGAAAGTGCAGAAATACGTCGAAGACA ACCAT GT 360
Db 318	GCCCCTTCTCGGGAGACCCGTCGCCAGATGCACAAAACAAAGCAAGACMACCCATGT 377
OY 361	GGCGGGGGCAATGTCATCTATACCAGCACTCCCTCCATACATACGCGCTGTCTGTA AAC AC 420
Db 378	GTCACATGGATGTGCTCATATACCAGGAAGCACCCCTTACTACGCTGTGCTGCAATA AT AC 437
OY 421	CCTTACACAGATCCCACTGCTCCCAAGTGGTTCCTGTATGCAAGGCCAAACCCCTGC AG 480
Db 438	CCTTACACAGGACCACTGCTCCCAAAGTGCCTTCGCGATGCAAGGCC-AAACCCCTGC AG 496
OY 481	AATGGGGGACTGCTGCTCGGGCAATTAAGGGSAAATGCAAGTTCACTGTGCTCGTCCG AC 540
Db 497	AATGGGGGAGTGTGTCGCCACACACAGGAAATCCAGTTTACCTGTGCTGTGCTCA GA C 556
OY 541	CAGTTCAAGGGGAAATCTGTGAATATAGTTCTGATGACTGTATGTTGGCATGGCTAC 600
Db 557	CAGTATTAAGGGGAAATCTGTGAATATAGTTCGCGACACAGCTTATGTCGATGATGCC TAC 616
OY 601	TCTTACCGAGGAAAAATGAAT--AGGCACTACACAGCATAGCGGTCTTACTGGA ACT C 659
Db 617	TCTTACCGAGGCAAAATGAGTAAAGACATCAACCAACCATGCTTTACTGSA MAC TC 676
OY 660	CCAC 663

```

Db      677 CCAC 680
|||||
RESULT 13
LOCUS   BF785781
DEFINITION 929 bp mRNA linear EST 12-JAN-2001
602112402F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:4240465
5', mRNA sequence.
ACCESSION BF785781
VERSION   BF785781.1 GI:12090817
KEYWORDS EST.
SOURCE   house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 929)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHM9855 row: d column: 02
High quality sequence stop: 625.
Location/Qualifiers
1. 929
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_1db="IMAGE:4240465"
/clone_1lb="NCI_CGAP_Kid14"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: kidney; Vector: pCMV-SPORT6; Site: 1; NotI;
Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library. |"

BASE COUNT 238 a 269 c 237 g 185 t
ORIGIN
Query Match 22.9%; Score 384.6; DB 12; Length 929;
Best Local Similarity 79.8%; Pred. No. 2.5e-98;
Matches 478; Conservative 0; Mismatches 119; Indels 2; Gaps 2;

QY 372 ATGTCATTTACCCAGAGTCCCTCTACTACCGCTGTCTGTAACACCTTTACACAG 431
||| ||||||||||||||| ||||||||||||||| ||||||||||||||| ||
Db 19 ATGCTCATTTACCCAGAGTCCCTCTACTACCGCTGTCTGTAACACCTTTACACAG 78
QY 432 TCCACAGTCTCCCAAGTGTCTCTGTATGAGAGCCAAACCCCTGCAGATGGGGCTAC 491
||| ||||||||||||||| ||||||||||||||| ||||||||||||||| ||
Db 79 ACCAGACTGTCCCAAGTGTCTCTGTATGAGAGCCAAACCCCTGCAGATGGGGAGT 138
QY 492 CTGCTCCCGGATTAAGGAGATCCCAAGTTCACCTGTGCTGCTCCGACCAAGTTAAAGG 551
||| ||||||||||||||| ||||||||||||||| ||||||||||||||| ||
Db 139 CTGTCCCGGACAGAGAGATCCCAAGTTCACCTGTGCTGCTCCGACCAAGTTAAAGG 198
QY 552 GAAATTTGTGAAATAGTTCATGATGCTATGTTGGCATGGCTACTTTACCGAG 611
||||| ||||||||||||||| ||||||||||||||| ||||||||||||||| ||
Db 199 GAAATTTGTGAAATAGTTCATGATGCTATGTTGGCATGGCTACTTTACCGAGG 258
QY 612 GAAATGTAATGAGACATCAACAGATGCTGCTTTACTGAACTCCCACTCTCTCTT 671
||||| ||||||||||||||| ||||||||||||||| ||||||||||||||| ||
Db 259 GAAATGTAATGAGACATCAACAGATGCTGCTTTACTGAACTCCCACTCTCTCTT 318
QY 672 GCAGAGATTTACACATGTTATGAGAGATGTAAGAACCATGATGGGAGACAA 731
||||| ||||||||||||||| ||||||||||||||| ||||||||||||||| ||
Db 319 GCAGAGATTTACACATGTTATGAGAGATGTAAGAACCATGATGGGAGACAA 378

```

```

QY 732 TTTCTGAGAAACCCAGATCGGACGAAAGCCCTGTCTTTATTAAATTACCAATGA 791
||||| ||||||||||||||| ||| ||||||||||| ||| ||| ||| |||
Db 379 CTTCGTGAGAAACCCAGATGAGACACAAACCTGTCTTTCTGTCAGAGTGAACAGTGA 438
QY 792 CAAGGTGAATGGGAATACCTGTGATGCTCAGCTGTGACGCCAGGAGCTTGCTACCC 851
||||| ||||||||||||||| ||||||||||||||| ||||||||||||||| |||
Db 439 GAAGGTGAATGGGAATACCTGTGATGCTCAGCTGTGACGCCAGGAGCTTGACCC 498
QY 852 AGAGGAAGCCCACTGAGCCATCAACCAAGCTTCCGGGGTTTGACTCTGTGGAAGAC 911
||| ||||||| ||||| ||| ||| ||| ||| ||| ||| ||| |||
Db 499 AGTGAAGAGCTTCTGTGAGGCTGTGATGAGAGC-TGCAGGCTTCGAGCTCCGGGAAGC 557
QY 912 TGAGATGACAGAGAGAAAGATCAAGATCTATGAGGCTTTAAGACACAGCGGGCA 970
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 558 GAGAGTACGTGAACACGACGATCAAGCTATCTACGG-GCCTTTAAGACACAGGAGCA 615

RESULT 14
LOCUS   BF384535
DEFINITION 969 bp mRNA linear EST 27-NOV-2000
602046804F1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:4196376 5',
mRNA sequence.
ACCESSION BF384535
VERSION   BF384535.1 GI:11365840
KEYWORDS EST.
SOURCE   house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 969)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHM9533 row: g column: 01
High quality sequence stop: 692.
Location/Qualifiers
1. 969
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_1db="IMAGE:4196376"
/clone_1lb="NCI_CGAP_L19"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: liver; Vector: pCMV-SPORT6; Site: 1; NotI;
Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.9 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library. |"

BASE COUNT 258 a 278 c 260 g 172 t 1 others
ORIGIN
Query Match 22.1%; Score 372.4; DB 12; Length 969;
Best Local Similarity 76.5%; Pred. No. 7.9e-95;
Matches 484; Conservative 0; Mismatches 142; Indels 7; Gaps 2;

QY 1 ATGTTTCCAGAGATGTCATGATCTGATGCTGTTAATGAGCTGTGGGAAGACA 60
||||| ||||||||||||||| ||| ||||||||||| ||| ||| ||| |||
Db 64 ATATTTGTAGAGATGTTGGTTCCTGCTGCTGCTGCTAATGCGCTGGGGGAAGTCA 123
QY 61 GCCTGTGGTTCCTCCATGATGCTTTATTTGAAAGCCGAGACCCAGACTGACCCCTGAC 120
||||| ||||||||||||||| ||| ||||||||||| ||| ||| ||| |||
Db 124 GTATGTTGGCTCTGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 183
QY 121 CAGTATGATTACAGCTACGAGGATTTATATCAGAGAGAAACACCAAGTACACATTACC 180

```


Db 184 GACTATTACTACAGCTATGACAGCTCCAGACCCGAGGAGACCCAGTGTACACCAACC 243
OY 181 CAGTCGAGAAATCTGACTGTACTACATGAGACCAAGCTGATCCATGCCAGCCCAAC 240
Db 244 ACCCCGAGAAACCCGAGTGTACT-----ATGAGAGCATGATCCATGCCAGTCCAAC 297
OY 241 CCTGTGAACAGGTGGGAGCTGCTGCATGAGGACACTTCACATGACAGTGTG 300
Db 298 CCTGTGAACAGGTGGGAGCTGCTGCATGAGGAGACTTCATGATGAGTGTG 300
OY 301 GCTCTCTTCTGCGAATAGTGTGCAAAAGTCAAAATACGTGCAAGACACCATGT 360
Db 358 GCCCCTCTCGGGAGCGGTGCCAGATGCAAAACAGTGCAGACCAACCATGT 417
OY 361 GGGCGGGGCGCATGTCTCATACCAAGTCTCCCTACTACCGCTGTGTCTGTAACAC 420
Db 418 GTCCATGTGTGATGCTCTATACCCAGACACCCCTACTACCGCTGTGTCTGTAACAC 477
OY 421 CTTTACACAGTCCAGCTGCTCCAAAGTGTCTGTATGACAGGCAACCCCTGCAG 480
Db 478 CTTTACACAGGACGACGACTGCTCCAAAGTGTCTGCGCATGACAGGCAACCCCTGCAG 537
OY 481 AATGGGCTACCTGCTCCCGGATACGGAGATCCAGATTCACCTGTGCTGTCCAGAC 540
Db 538 AATGGGAGTGTCTGTCCCGACACAGACGAGATCCAGATTCACCTGTGCTGTCCAGAC 597
OY 541 CAGTTACAGGGAATTTCTGAAATAGTGTGTGATGACTGCTATGTGGCGATGCTAC 600
Db 598 CAGTATTAAGGGAATTTCTGAAATAGTGTGTGCGGACGACTTATGTGTGATGCTAC 657
OY 601 TCTTACCGAGGGAATGATAGACAGTCAAC 633
Db 658 TCTTACCGA-GGCAAGTGTATAGACAGTCAAC 689

RESULT 15
AM475402 654 bp mRNA linear EST 24-FEB-2000
LOCUS
DEFINITION
IMAGE:2503549.5' similar to TR:Q14520 Q14520 HGF ACTIVATOR LIKE
PROTEIN. ; mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 654)
Marr, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wyllie, T.,
Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, Y., Person,
B., Waller, J., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Rittler,
E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R. and Wilson, R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)
Contact: Marra M/Mashu-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:1018401
Seq primer: custom primer used
High quality sequence stop: 482.
Location/Qualifiers
1..654
/organism="Mus musculus"
/strain="C57BL"
/db_xref="taxon:10090"

/clone="IMAGE:2503549"
/clone_lib="Sugano mouse kidney mklA"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: Kidney; Vector: pME18S-FL3; Site_1: DraIII
(CACTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCTTTTTTTTTTTTTTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor (TGTGGCTTGGG), digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCTGTCTTAAAGCTGCG and 3' end
primer CGACTGTGAGCTGAGGACA."
BASE COUNT 156 a 193 c 174 g 131 t
ORIGIN

Query Match 22.0%; Score 370.6; DB 10; Length 654;
Best Local Similarity 74.5%; Pred. No. 2,1e-94;
Matches 482; Conservative 0; Mismatches 159; Indels 6; Gaps 1;

OY 1 ATGTTGCCAGAGTGTGATCTTCATGTTCTGCTGTTAATGCTGTGGTGGGAAAGACA 60
Db 13 ATATTGTGACAGATGTGGTGTCTGCTGTGCTGTATGCTGTGGTGGGAAAGTCA 72
OY 61 GCCTGTGGTCTCTCCGATGCTTATTATGGAACCCGAGACCCAGATCCCTGAC 120
Db 73 GTCAATTGCTCTCAGTATGATCTCTTCAATTTGCGCCCGACCAAGATGAGACCCCAT 132
OY 121 CAGTATGATACACTGACGAGGATTTATTCAGAAAGAACCCAGTATGACACTTAC 180
Db 133 GACTATTACTACAGCTATGAGCAGTCCAGCCAGACGAGACCCCATGTGACGACGAC 192
OY 181 CATGCTGAGAAATCTGACTGTACTACACTGAGACCAAGCTGATCCAGCCCAAC 240
Db 193 ACCCCTGTGAGAACCCCGACTGTACT-----ATGAGAGCATGATCATCCAGGCAAC 246
OY 241 CCTGTGAAACAGGTGGGAGTGCCTGTGTCATGAGGAGCACCCTTCACATGACGCTGCTG 300
Db 247 CCTGTGAAACAGGTGGGAGTGCCTGTGTCATGAGGAGGATGACTTGTGACGCTGCTG 306
OY 301 GCTCTTCTGTGGGAATTAAGTGCAGAAAGTCAAAATACGTGCAAGACACCATGT 360
Db 307 GCCCCTTCTGCGGAGCCGCTGTCAGACTGCACAAGACAGTGCAGACACCATGT 366
OY 361 GGGCGGGGCGAATGCTCATTTACCCAGAGTCTCCCTACTACCGCTGTGTCTGTAACAC 420
Db 367 GTCAATGTGATGCTCTATACCCAGAAACCCCTTACTACGCTGTGCTGCTGACATGAC 426
OY 421 CCTTACAGAGTCCAGCTCTCCCAAGTGTCTGTATGACAGGCCCAACCCCTGCAG 480
Db 427 CCTTACAGGAGACAGACTCTCCAAAGTGTGTGCGGATCAGAGCCCAACCCCTGCAG 486
OY 481 AATGGGCTACCTGCTCCCGGATTAAGGAGAGATCAAGTTCAGCTGTGCTGTGCTGCG 540
Db 487 AATGGGAGTGTGTTCCCGACACACAGCGAATCCAAAGGTTACTTGTGCTGTGCTGAC 546
OY 541 CAGTTCAAGGGGAAATTTCTGGAATAGTGTGATGACTCTATGTTGGCGATGCTAC 600
Db 547 CAGGTATAGGGAATTTCTGTGACATTAAGTGTGCGGAGTGTGTGCGTGTGCTGAC 606
OY 601 TCTTACCGAGGAAATGATAGACAGTCAACGATGCTGTGCTGTGCTGTGCTGTGCT 647
Db 607 TCTTACCGACCAAGGAGTGAAGACAGGCAAGCAAGCATGTCT 653

Search completed: March 6, 2003, 19:37:34
Job time: 2184.5 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 15:27:01 ; Search time 65.5 Seconds

(Without alignments)
7879.954 Million cell updates/sec

Title: US-09-912-559-2

Perfect score: 1683

Sequence: 1 atgttgcagagatgtctga.....aaagtgaagtgcctctaa 1683

Scoring table: IDENTITY_NUC

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued_Patents_NA.*
2: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	114.4	6.8	2033	1	US-08-148-910-14
2	114.4	6.8	2033	1	US-08-148-910-14
3	93	5.5	970	1	US-08-448-937A-14
4	93	5.5	970	1	US-08-148-910-3
5	72.2	4.3	329	1	US-08-448-937A-3
6	72.2	4.3	329	1	US-08-148-910-2
7	72.2	4.3	329	1	US-08-148-910-13
8	72.2	4.3	329	1	US-08-448-937A-2
9	72.2	4.3	1065	1	US-08-427-640-1
10	72.2	4.3	1065	1	US-08-427-640-5
11	72.2	4.3	1068	1	US-08-427-640-3
12	72.2	4.3	1137	4	US-09-553-498-9
13	72.2	4.3	1137	4	US-09-618-869-9
14	72.2	4.3	1314	2	US-08-811-949-48
15	72.2	4.3	1955	2	US-08-885-795A-39
16	72.2	4.3	2457	6	5344773-1
17	72.2	4.3	7360	1	US-08-286-740-1
18	72.2	4.3	7360	5	PCF-US95-09576-1
19	70.6	4.2	1068	1	US-08-811-949-60
20	70.6	4.2	1068	1	US-08-137-116-2
21	70.6	4.2	1068	2	US-08-427-640-7
22	70.6	4.2	1068	2	US-08-811-949-44
23	70.6	4.2	1068	2	US-08-811-949-46
24	70.6	4.2	1068	2	US-08-811-949-52
25	70.6	4.2	1068	2	US-08-811-949-58
26	70.6	4.2	1068	6	5223256-3
27	70.6	4.2	1163	2	US-08-558-269-5

28	70.6	4.2	1163	4	US-09-410-882-5	Sequence 5, Appl1
29	70.6	4.2	1170	2	US-08-811-949-64	Sequence 64, Appl1
30	70.6	4.2	1170	2	US-08-811-949-66	Sequence 66, Appl1
31	70.6	4.2	1314	2	US-08-811-949-50	Sequence 50, Appl1
32	70.6	4.2	1314	2	US-08-811-949-54	Sequence 54, Appl1
33	70.6	4.2	1314	2	US-08-811-949-56	Sequence 56, Appl1
34	70.6	4.2	1419	2	US-08-811-949-62	Sequence 62, Appl1
35	70.6	4.2	1738	6	5200340-1	Patent No. 5200340
36	70.6	4.2	1848	3	US-08-814-412-10	Sequence 10, Appl1
37	70.6	4.2	1974	2	US-08-811-949-38	Sequence 38, Appl1
38	70.6	4.2	2101	2	US-08-811-949-42	Sequence 42, Appl1
39	70.6	4.2	2162	1	US-08-119-512-3	Sequence 3, Appl1
40	70.6	4.2	2162	1	US-08-488-015B-3	Sequence 25, Appl1
41	70.6	4.2	2162	1	US-08-488-015B-25	Patent No. 5200340
42	67.8	4.0	1724	6	5200340-5	Patent No. 5200340
43	67.8	4.0	2497	6	5185259-2	Patent No. 5185259
44	66.4	3.9	2544	4	US-09-518-046-3	Sequence 3, Appl1
45	66	3.9	1233	1	US-08-254-922-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-08-148-910-14
Sequence 14, Application US/08148910
Patent No. 546593
GENERAL INFORMATION:
APPLICANT: Takeshi SHIMOMURA et al.
TITLE OF INVENTION: NO. 546593el Protein and Gene Encoding Said Protein
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
City: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch,
MEDIUM TYPE: 500 Kb Storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/148,910
FILING DATE: No. 546593ember 5, 1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX: 202-371-8856
TELEX:
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 2033 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: human
IMMEDIATE SOURCE:
LIBRARY: Pre-made Lambda phage Library,
LIBRARY: human liver(49, male) cDNA Library (Stratagene)
US-08-148-910-14

Query Match	Similarity	Score	DB	Length
Best Local	49.6%	114.4	DB 1	2033
Matches	605	Conservative	0	Mismatches 561; Indels 54; Gaps 10;
QY	469	AAACCCGTCAGATGAGGGCTACCTCTCCGGCCATTAAGCGGAGATCCAAAGTCAACCTGT	528	
DB	742	AAACCCGTCAGATGAGGGCTACCTCTCCGGCCATTAAGCGGAGATCCAAAGTCAACCTGT	528	
QY	529	GCCATGCCGACCACTTTCAGGGGAATTCGTGAATAGGTTCTGATGA---CTGGCAT	585	
DB	802	GCCATGCCGACCACTTTCAGGGGAATTCGTGAATAGGTTCTGATGA---CTGGCAT	585	
QY	586	GTTGGCGATGCTACTCTTAACCAAGGGAATGAATAGAGAGTCAACCAAGCATCGCTGC	645	
DB	862	TTCGGGGAAGCGCACTGGGTACCGTGGCGGAGAGAGTCAACCAAGCATCGCTGC	645	
QY	646	CTTTACTGGAAATCCCACTTCCTTTGCGAGGAATTAACAACATGTTTATGAGAGATGCT	705	
DB	922	CTGGCCCTGGAAATCCCACTTCCTTTGCGAGGAATTAACAACATGTTTATGAGAGATGCT	705	
QY	706	GAAACCCATGGATTTGGGGAACAAATTTCTGCAAAAACCCAGATCGGAGGAAAGCCC	765	
DB	982	GCCCTCTGGGCTGGGGCCCATGGCTCTTGGCGGATCCGAGCAATGAGAGAGGCCC	1041	
QY	766	TGCTGTTTAAATTAAGTTACCAATGACAGGTAAGTAATGGAATACGTGATGTCTCAGCC	825	
DB	1042	TGCTGTTTAAATTAAGTTACCAATGACAGGTAAGTAATGGAATACGTGATGTCTCAGCC	825	
QY	826	TGCTCAAGCCCAAGAGATGTTGCTTACCCAGAGGAAGCCCCTAGCATCAACCAAGCTT	885	
DB	1099	TGCTCAAGCCCAAGAGATGTTGCTTACCCAGAGGAAGCCCCTAGCATCAACCAAGCTT	885	
QY	886	CCGAGGTTTGAATCTCTGAGAAAGATGAGATAGAGAGAGAGATGATCAAGAGATATAT	945	
DB	1150	CCGAGGTTTGAATCTCTGAGAAAGATGAGATAGAGAGAGAGATGATCAAGAGATATAT	945	
QY	946	GGAGGTTTAAAGAGAGCGGGGCAAGCAACCATGAGAGCGTCTCATGCTCTGCTGT	1005	
DB	1210	CTCGGGCCAGCATATATCGGGGCTCTCTCTGCGCCGGCTCGAGACCCCTG-----	1262	
QY	1006	CCCTGTACCATTCATGCCCCAGGGGCCACTTCTGTGTGGGGCGCTGATCCACCCCTGC	1065	
DB	1263	CCCTGTACCATTCATGCCCCAGGGGCCACTTCTGTGTGGGGCGCTGATCCACCCCTGC	1065	
QY	1066	TGGGTCTCTACGCGCCGCACTGACACCGCAACATTAATAAACGAGATCTAA-----AGGTG	1119	
DB	1321	TGGGTCTCTACGCGCCGCACTGACACCGCAACATTAATAAACGAGATCTAA-----AGGTG	1119	
QY	1120	GTCCTAGGGGGAACAGACCTGAAAGAAAGAAATTTATGAGAGAGATTTAGGGTTCAG	1179	
DB	1381	GTCCTAGGGGGAACAGACCTGAAAGAAAGAAATTTATGAGAGAGATTTAGGGTTCAG	1179	
QY	1180	AAATATTTCAAGTACACGCACTCATGTAAGAGATGATGATTTCCCAATAGATATGCA	1239	
DB	1441	AAATATTTCAAGTACACGCACTCATGTAAGAGATGATGATTTCCCAATAGATATGCA	1239	
QY	1240	TTCCTCAAGTTAAAGCCAGTGTATGTCATCTGTCTAGATCAAAATACGTGAAGACT	1299	
DB	1498	CTGATCCGCTGAAGAAAGAAAGGGAACGCGTGTGTCACACGCTCGCATGTTGTCAACCC	1557	
QY	1300	GTTGTGTTTCCCTGATG-----GATCTTTTCCCTGTGGAGAGAGTGCACATCTCGGC	1353	
DB	1558	ATCTGCTGCTCCGAGCCGAGCAACATCTTCCCGCAGAGACACAAGATGCTCAATTTGCGGCG	1617	
QY	1354	TGGGG-----TGTTCAGAAAGAGAAAGGATGCCGCAAGCTCTGTATGCCAAATG	1407	
DB	1618	TGGGG-----TGTTCAGAAAGAGAAAGGATGCCGCAAGCTCTGTATGCCAAATG	1407	
QY	1408	AAAGCTGATGAGCAACATTTGTGCAATCCCGGCACTCTATAGACACATGATGATGAC	1467	
DB	1678	AAAGCTGATGAGCAACATTTGTGCAATCCCGGCACTCTATAGACACATGATGATGAC	1467	
QY	1468	AGTATGATCTGTGAGGAAATCTTCAAGAAACCTGGGCAAGACACCTGCCAGGATGACT	1527	

```

Db      1738  AACATGCTCTGTGGCGCGGCTACTT---GCACTGCAGATCCGACGCTCCAGGGGGACATCA 1794
Oy      1528  GAGAGGCCCCCTGACTGTGAGAGAGAGCGCACCCTACTACGCTCTATGGGATAGTACGCTGG 1587
Db      1795  GGGGGGGCCCCCTGGCTCGAGAGAAACGGCGTGGCTTACCTCTAAGGCAATCATCAGCTGG 1854
Oy      1588  GGCGCTGAGAGTGTGAG-----AGAGGCGCAGGGGCTGTACACCCAAAGTTACCAAAATTCCTG 1641
Db      1855  GGTGACGCGCTCGGGCGGCTCCACAGACGGGGGTCTACACCCCGGTGGCAACATATGTG 1914
Oy      1642  AATTGATCAAGCCACCAT 1661
Db      1915  GACTGATCAACGACCGGAT 1934

RESULT 2
US-08-448-937A-14
Sequence 14, Application US/08448937A
Patent No. 5677164
GENERAL INFORMATION:
APPLICANT: Takeshi SHIMOMURA et al.
TITLE OF INVENTION: No. 5677164el Protein and Gene Encoding Said Protein
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch,
MEDIUM TYPE: 500 Kb Storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/448,937A
FILING DATE: May 24, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/148,910
FILING DATE: No. 5677164ember 5, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX: 202-371-8856
TELEX:
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 2033 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: human
IMMEDIATE SOURCE:
LIBRARY: Pre-made lambda phage library,
LIBRARY: human liver(49, male) cDNA Library (Stratagene)
US-08-448-937A-14
Query Match 6.8%; Score 114.4; DB 1; Length 2033;
Best Local Similarity 49.6%; Pied. No. 3.5e-24;
Matches 605; Conservative 0; Mismatches 561; Indels 54; Gaps 10.
Oy      469  AACCCCTGCCAGAAATGGGCGCTACCTGCTCCGGCATAGCGGAGATCCAAAGTCACTGT 528
Db      742  AGCCCTTCCTGGAAGGGGCGACCTGCACTGATCGTGGCCACCGGAGCACCGTGTGT 801

```

```

OY 529 GCGTCCCGACAGCTTCAGAGGGAATTCCTGTAATAGTTCATGATGA---CTGCTAT 585
    ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 802 GCGTCCCGACAGGCTCTGAGGAGCGGCTCGACACATCGAGCTGTGATGACGCTGCTTC 861
OY 586 GTTGGCGATGGCTACTCTTACCGAGGGAATAATAGACATGACATCAACAGCATGCTGC 645
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 862 TTGGGGAGACGGCATGGGTACCGGTGGGGTGGCCAGCAGCCTCAGGCTCGGGGCTCAGCTGC 921
OY 646 CTTTACTGGAATCCCACTCTCTCTTGAGAGAGATTTACAACTGTTATGAGAGATGCT 705
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 922 CTGGCCCTGGAACTCGATGCTGTACAGAGAGCTGACGTGAGACTCGGTGGGCGCGG 981
OY 706 GAAACCATGGGATTGGGGACACAAATTTCTGCAAGAAACCCAGATGCGAGAAAGCCC 765
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 982 GCGCTGTGGGGCTGGGGCCCGGATGCTGCTGCGGAGATCGGAGATGCGAGAGAGGCC 1041
OY 766 TGGTCTTTATTAAGTTAACATGACAGGTGAATGGGAATGCTGTGATGCTCAGCC 825
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1042 TGGTGTCTACGT---GGTGAAGAGACAGCGCGCTCTCCGAGATGACTGCGCCGAGAGCC 1098
OY 826 TGCTCAGCCAGGAGCTTGCTCTACAGAGGAAAGCCCACTGAGCATCAACCAAGCTT 885
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1099 TCGCAATCCCT-----CACAGAGTCCAACTGTCAACCGGATCTCTGGCAGACCTG 1149
OY 886 CCGGGGTTTACTCTGTGTGAAGACTGAGATGACAGAGAGAAATCAAGAGATCTAT 945
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1150 CTTGAGCCAGCCTCCCGGGGCGCCAGGCTGCGGAGAGACAAAGAGAGAGCTTC 1209
OY 946 GGAGGCTTTAAGACAGCGGCGGCAAGCACCATGCGAGCGCTCCAGTCCGCTG 1005
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1210 CTGGGCGACAGTATCGGGGCTCTCTCGCTGCGCGGCTCGACCCCTG----- 1262
OY 1006 CCTGTGACATCTCATGCCCCAGGGCCACTTCTGTGTGGGGCGTGTATTCACCCCTGC 1065
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1263 --GCTGGCCCGCATCTAATCGGGAGACGCTTCTGCGCGGGAGCTGGTCCACACCTGC 1320
OY 1066 TGGGTGCTCACTGCTGCACTGCACGACGACATMAAAACAGACATCTAA-----AGGTC 1119
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1321 TGGGTGCTGTGCGCGCCGCACTGCTTCCACAGCCCCCGAGGAGACGCTCTCCCTG 1380
OY 1120 GTGCTAGGGAGACAGACCTGTGAAGAAAGAAATTTATGAGAGAGCTTTAGGTTGAG 1179
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1381 GTGCTGGGGCAGACTTCTTCAACCGCAGCAGGACGACGACGCTTGGGCACTGAG 1440
OY 1180 AAGATATTCAAGTACAGCCACTCAATGAAGAGATGATCCCAATGATATTGCA 1239
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1441 AAGTACATCCCGTACACCCCTGATCGGTTCATMAACCCAG---CGACACAGACCTGTC 1497
OY 1240 TTGCTCAAGTTAAGCATGATGATGTCATGCTCTTCAATTCAAATACCTGAAGACT 1259
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1498 CTGATCCGGCTGAAGAAAGGAGGACGCTGTGCACACGCTGCGAGTTCGTCAGGCC 1557
OY 1300 GTGTGCTTGGCTGATG-----GATCCTTTCCTGAGAGTGGCCACATCTCTGCGC 1353
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1558 ATGTGCTGCGCCAGCCCGGACAGCACTTCCCGAGAGACAAAGTGCAGATTCGCGGC 1617
OY 1354 TGGGG-----TGTACAGAAACAGAAAGAGTCCCGGACGCTCTGATGCTCCAAATC 1407
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1618 TGGGGCCTCTTGTGATGAACGTGAGCGGCTACGCTCCCTGCGGAGAGGCCCTGCTG 1677
OY 1408 AAGCTGATTGCAACACTTTGTGCAACTCCCGCACTCTATGACCACTGATTTGATGAC 1467
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1678 CCGCTGGGCGCGACCAAGTGCAGCAGCCCTGAGGCTTACGAGCGCGACATCAGCCCC 1737
OY 1468 AGTATGATCTGTGAGAAATCTTCAGAAACCTGGGCAACACACCTGCGAGGAGTACTCT 1527
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1738 AACATGCTGTGTGCGGCTACTT---CGATGCAAGTCTGAGGCTGCGAGGAGAGCTCA 1794
OY 1528 GGAGGCGCCCTGACCTGTGAGAGAGCGCACACTACTAGCTTATGGATAGTGAAGCTGG 1587
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1795 GGGGGGCGCGCTGCGCTGCGAGAAAGCGCGTACTTACCTTACGAGCATCATCAGCTGG 1854

```

```

OY 1588 GGCGTGGAGTGTGAG-----AAGAGCGCAGGGGCTTACACCAAGTTACCAAAATTCCTG 1641
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1855 GGTGACGCGCTCGGGCGGCTCCACAAAGCCGGGGGTCTTACACCGCGGTGGCAACTATGTG 1914
OY 1642 AATTGATCAAAAGCCACCAT 1661
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1915 GACTGATCAACGACCGGAT 1934

```

```

RESULT 3
US-08-148-910-3
; Sequence 3, Application US/08148910
; Patent No. 546593
; GENERAL INFORMATION:
; APPLICANT: Takeshi SHIMOMURA et al.
; TITLE OF INVENTION: No. 546593el Protein and Gene Encoding Said Protein
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch,
; MEDIUM TYPE: 500 Kb Storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/148,910
; FILING DATE: No. 546593ember 5, 1993
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Marten M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX: 202-371-8856
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 970 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: human
; IMMEDIATE SOURCE:
; LIBRARY: Pre-made Lambda phage Library, human liver (49, male)
; LIBRARY: cDNA Library (Stratagene)
US-08-148-910-3

Query Match 5.58; Score 93; DB 1; Length 970;
Best Local Similarity 52.7%; Pred. No. 6; Le-18;
Matches 357; Conservative 0; Mismatches 290; Indels 30; Gaps 6;

OY 1009 CTACCATCTTCATATCCCGAGGCACTTCTGTGTGGGGCGCTGATCACCCTCGCTGG 1068
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 201 CTGGCCCGCATCTACATGTGGGAGCAGCTTCTGCGCGGGAGCGTGTGTCAACACTGTGTG 260
OY 1069 GTGTCTACGTGCTGCCCATCTGCACGACATATAAAGCAGACA-----TCTAAAGTGTG 1122
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 261 GTGTGTGTGGCGCGCCACTGTCTTCCACAGCCGCCCGAGGACAGCGTTCGCTGTGTG 320
OY 1123 CTAGGAGACAGACCTGAAGAAAGAAATTTCAATGACAGAGCTTTAGGCTGACAGAG 1182
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 321 CTGGGCGACAGCTTCTTCAACCGACGAGCAGTACGACGACCTTGGCATCGAAG 380
QY 1183 ATATTGAGTACGACGACCTACATGAAAGATGAGATTCCCAATGATATTGCTATG 1242
Db 381 TACATCCGCTACGACCTGATCGGTTCACCCCA---CGACACGACGACCTGCTCG 437
QY 1243 CTCAGTAAAGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1302
Db 438 ATCCGCTGAAGAAAGGAGGAGCGCTGTCACACGCTGCACTGCTGCTGCTGCTGCTG 497
QY 1303 TGCCTCCGATG-----GCTCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1356
Db 498 TGCCTCCGACCGCGACGACCTTCCCGACGACGACGACGACGACGACGACGACG 557
QY 1357 GG-----TGTACAGAAACAGAAAGAGGCTCCGACGCTCTGCTGCTGCTGCTG 1410
Db 558 GGCACCTGTGATGAAACGTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 617
QY 1411 CTGATTGCAACCTTTGTGCACTCCGCGCACTGCTGCTGCTGCTGCTGCTGCTGCTG 1470
Db 618 CTGCTGCGGACGACGACGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 677
QY 1471 ATGATCTGTGAGAAATCTTCAAGAACTGCGCAAGACCTGCTGCTGCTGCTGCTG 1530
Db 678 ATGCTCTGTGCGGCTACTT---CGACTGCAAGTCCGACGCTGCTGCTGCTGCTG 734
QY 1531 GGGCCCCGACCTGTGAGAGGAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1590
Db 735 GGGCCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 794
QY 1591 CTGAGTGTGAG-----AAGAGCGAGGGGTCTACACCAAGTTACCAATTTCTGAT 1644
Db 795 GACGGCTCGGGGCGCTCCACAAGCGGGGTCTACACCGCGGCAACTGTGTGAC 854
QY 1645 TGGATCAAGCCACCAT 1661
Db 855 TGGATCAAGCCGAT 871

RESULT 4
US-08-448-937A-3
; Sequence 3, Application US/08448937A
; Patent No. 5677164
; GENERAL INFORMATION:
; APPLICANT: Takeshi SHIMOMURA et al.
; TITLE OF INVENTION: No. 5677164el Protein and Gene Encoding Said Protein
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch,
; MEDIUM TYPE: 500 Kb Storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448, 937A
; FILING DATE: May 24, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/148, 910
; FILING DATE: No. 5677164ember 5, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850

TELEFAX: 202-371-8856
; TELTEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 970 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: human
; IMMEDIATE SOURCE:
; LIBRARY: Pre-made Lambda phage Library, human liver (49, male)
; LIBRARY: cDNA Library (Stratagene)
US-08-448-937A-3

Query Match 5.5%; Score 93; DB 1; Length 970;
Best Local Similarity 52.7%; Pred. No. 6.1e-18;
Matches 357; Conservative 0; Mismatches 290; Indels 30; Gaps 6;

QY 1009 CTGACATCTCCATGCCCCAGGCGCACTTCTGTGTGGGCGCTGATCCACCTGCTG 1068
Db 201 CTGGCCGCCATCTACATGCGGGACAGCTTCTGGCGCGGAGGCTGTCCACACCTGCTG 260
QY 1069 GTGCTACGTCTGCCACTCTGACCGACATATAAACGACA-----TCTAAGGTGTG 1122
Db 261 GTGCTGTGGCGCGCCACTGCTTCTCCACAGCCCCCGAGGACAGCGTCTCCGTGTG 320
QY 1123 CTAGGGACACGACCTGTAAGAAAGAAATTTCTATGACGACGCTTTAGGGTGCAGAG 1182
Db 321 CTGGGCGACACTTCTTCACCGCAGCAGGAGCTGACGACACCTTGGCATGGAAG 380
QY 1183 ATATTGAGTACGACGACCTACATGAAAGATGAGATTCCCAATGATATTGCTATG 1242
Db 381 TACATCCGCTACGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 437
QY 1243 CTCAGTAAAGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1302
Db 438 ATCCGCTGAAGAAAGGAGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 497
QY 1303 TGCCTCCGATG-----GCTCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1356
Db 498 TGCCTCCGACCGCGGACGACCTTCCCGACGAGACACAACTGCTGCTGCTGCTGCTG 557
QY 1357 GG-----TGTACAGAAACAGAAAGGAGGCTCCGACGCTCTGATGCCAAAGTCAAG 1410
Db 558 GGCACCTGTGATGAAAGCTGAGCGGCTACTCCAGCTGCTGCTGCTGCTGCTGCTGCTG 617
QY 1411 CTGATTGCAACCTTTGTGCACTCCGCGCACTGCTGCTGCTGCTGCTGCTGCTGCTG 1470
Db 618 CTGCTGCGGACGACGACGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 677
QY 1471 ATGATCTGTGAGAAATCTTCAAGAACTGCGCAAGACCTGCGAGGCTGACTGTGA 1530
Db 678 ATGCTCTGTGCGGCTACTT---CGACTGCAAGTCCGACGCTGCGAGGGAGCTGAGG 734
QY 1531 GGGCCCCGACCTGTGAGAGGAGGAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1590
Db 735 GGGCCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 794
QY 1591 CTGAGTGTGAG-----AAGAGCGAGGGGTCTACACCAAGTTACCAATTTCTGAT 1644
Db 795 GACGGCTCGGGGCGCTCCACAAGCGGGGTCTACACCGCGGCAACTGTGTGAC 854
QY 1645 TGGATCAAGCCACCAT 1661
Db 855 TGGATCAAGCCGAT 871

RESULT 5
US-08-148-910-2
; Sequence 2, Application US/08148910
; Patent No. 546593

```

: GENERAL INFORMATION:
: APPLICANT: Takeshi SHIMOMURA et al.
: TITLE OF INVENTION: No. 546593el Protein and Gene Encoding Said Protein
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Wenderoth, Lind & Ponack
: STREET: 805 Fifteenth Street, N.W., #700
: CITY: Washington
: STATE: D.C.
: COUNTRY: U.S.A.
: ZIP: 20005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 5.25 inch,
: MEDIUM TYPE: 500 Kb Storage
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: MS-DOS
: SOFTWARE: Wordperfect
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/148,910
: FILING DATE: No. 546593ember 5, 1993
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Warren M. Cheek, Jr.
: REGISTRATION NUMBER: 33,367
: REFERENCE/DOCKET NUMBER:
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-371-8850
: TELEFAX: 202-371-8856
: TELEX:
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 329 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: ORIGINAL SOURCE:
: ORGANISM: human
: IMMEDIATE SOURCE:
: LIBRARY: Quick-cloneTM human liver cDNA (Clontech)
: US-08-148-910-2

Query Match          4.3%; Score 72.2; DB 1; Length 329;
Best Local Similarity 59.3%; Pred. No. 5,7e-12;
Matches 163; Conservative 0; Mismatches 103; Indels 9; Gaps 2;

QY 1381 TCCCGCCAGCTCCTGGATGCCAAGTCAAGCTGATTGCCAACCTTTGTGCMACTCCCGC 1440
DB 52 TCCAGCTCCCTCGCGGAGGCGCTGTCCTCCCTGTCGCCGACCAACAGTGCAGCGCCT 111
QY 1441 CAACCTATGACACATGATGTGATGATGATGTCGAGGAATCTTCGAACCT 1500
DB 112 GAGGTCTACGGCGCGGACATCAAGCCCAACATGCTGTGCGGCTACTT---CGACTGC 168
QY 1501 GGGCAAGACACTGCCAGGCTGACTGTGAGAGCCCTCGACTGTGAGAAGACGGCACC 1560
DB 169 AAGTCGACGCTCCACAGGGGAGCTAGAGGGGAGCCCTGCTGCGAGGAAGACGGCCTG 228
QY 1561 TACTACGCTATGAGATAGTGAAGCTGGGGCTCGAGAGTGTGAG-----AAGAGCCAGGG 1614
DB 229 GCTTACCTCTACGGCATCATCACTGAGGCTGACGGCTGGGGGCTCCCAAGCGGGG 288
QY 1615 GTCTACACCCAGTTACCAAAATTCCTGAATTGGAT 1649
DB 289 GTCTACACCCGCGTGGCAACTATGTGAGCTGGAT 323

RESULT 6
US-08-148-910-13
; Sequence 13, Application US/08148910
```

```

: Patent No. 546593
: GENERAL INFORMATION:
: APPLICANT: Takeshi SHIMOMURA et al.
: TITLE OF INVENTION: No. 546593el Protein and Gene Encoding Said Protein
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Wenderoth, Lind & Ponack
: STREET: 805 Fifteenth Street, N.W., #700
: CITY: Washington
: STATE: D.C.
: COUNTRY: U.S.A.
: ZIP: 20005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 5.25 inch,
: MEDIUM TYPE: 500 Kb Storage
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: MS-DOS
: SOFTWARE: Wordperfect
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/148,910
: FILING DATE: No. 546593ember 5, 1993
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Warren M. Cheek, Jr.
: REGISTRATION NUMBER: 33,367
: REFERENCE/DOCKET NUMBER:
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-371-8850
: TELEFAX: 202-371-8856
: TELEX:
: INFORMATION FOR SEQ ID NO: 13:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 329 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: ORIGINAL SOURCE:
: ORGANISM: human
: IMMEDIATE SOURCE:
: LIBRARY: Quick-cloneTM human liver cDNA (Clontech)
: US-08-148-910-13

Query Match          4.3%; Score 72.2; DB 1; Length 329;
Best Local Similarity 59.3%; Pred. No. 5,7e-12;
Matches 163; Conservative 0; Mismatches 103; Indels 9; Gaps 2;

QY 1381 TCCCGCCAGCTCCTGGATGCCAAGTCAAGCTGATTGCCAACCTTTGTGCMACTCCCGC 1440
DB 52 TCCAGCTCCCTCGCGGAGGCGCTGTCCTCCCTGTCGCCGACCAACAGTGCAGCGCCT 111
QY 1441 CAACCTATGACACATGATGTGATGATGATGTCGAGGAATCTTCGAACCT 1500
DB 112 GAGGTCTACGGCGCGGACATCAAGCCCAACATGCTGTGCGGCTACTT---CGACTGC 168
QY 1501 GGGCAAGACACTGCCAGGCTGACTGTGAGAGCCCTCGACTGTGAGAAGACGGCACC 1560
DB 169 AAGTCGACGCTCCACAGGGGAGCTAGAGGGGAGCCCTGCTGCGAGGAAGACGGCCTG 228
QY 1561 TACTACGCTATGAGATAGTGAAGCTGGGGCTCGAGAGTGTGAG-----AAGAGCCAGGG 1614
DB 229 GCTTACCTCTACGGCATCATCACTGAGGCTGACGGCTGGGGGCTCCCAAGCGGGG 288
QY 1615 GTCTACACCCAGTTACCAAAATTCCTGAATTGGAT 1649
DB 289 GTCTACACCCGCGTGGCAACTATGTGAGCTGGAT 323

RESULT 7
US-08-148-937A-2
```

Sequence 2, Application US/08448937A
 Patent No. 5677164
 GENERAL INFORMATION:
 APPLICANT: Takeshi SHIMOMURA et al.
 TITLE OF INVENTION: No. 5677164e1 Protein and Gene Encoding Said Protein
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Wenderoth, Lind & Ponack
 STREET: 805 Fifteenth Street, N.W., #700
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 5.25 inch,
 MEDIUM TYPE: 500 Kb Storage
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: MS-DOS
 SOFTWARE: Wordperfect
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/448,937A
 FILING DATE: May 24, 1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/148,910
 FILING DATE: No. 5677164e1ember 5, 1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Warren M. Cheek, Jr.
 REGISTRATION NUMBER: 33,367
 REFERENCE/DOCKET NUMBER:
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-371-8850
 TELEFAX: 202-371-8856
 TELEX:
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 329 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 ORIGINAL SOURCE:
 ORGANISM: human
 IMMEDIATE SOURCE:
 LIBRARY: Quick-clone™ human liver cDNA (Clontech)
 (S-08-448-937A-2

Query Match Similarity	4.3%	Score 72.2	DB 1	Length 329
Best Local Similarity	59.38	Pred No. 5.7e-12		
Matches 163	Conservative 0	Mismatches 103	Indels 9	Gaps 2
QY	1381	TCGCCGCACTCTCTGGATGCGCAAAAGTCAAGCTGATTTGGCCAACTTTGTGCAACTCCGAC	1440	
Db	52	TCGAGCTCCCTTCGCGGAGGCCCTGTGCTCCCTGGTCCGCGCACCAAGTGCACAGCCTT	111	
QY	1441	CAACTCTATGACACCAATGATTTGATGACATATGATCTGTGAGGAAATCTTCGAAACTT	1500	
Db	112	GAGCTTCACGGGGCGCAGCATACGCCCAACAATGCTCTGTGCGGCTACTTT--CGACTGC	168	
QY	1501	GCGCAGACACCTGCCACGAGGTGACTGTGAGAGCCGCCCTGCACCTTGTGAGAAGCAGCACC	1560	
Db	169	AACTCGAGCGCTCGCAGGGGAGACTCAGGGGGGCCCTCGGCTCGCAGGAAGAAGAGCGCTG	228	
QY	1561	TACTACGTCTATGGATAGTGAAGCTGGGGCTCGAGTGTGAG-----AAGAGGCCAGGG	1614	
Db	229	GCTTACTCTATCGGCAATCATCACTACCTGGGTGTGACGGCTCGGCGGCTCCACAAGCGCGGG	288	
QY	1615	GTCACACCAAGTATACCAATTCCTCGAATTGGAT	1649	
Db	289	GTCATACCCGCGTGCGCAACTATGTGAGACTGAT	323	

US-08-448-937A-13
Sequence 13, Application US/08448937A
Patent No. 5677164
GENERAL INFORMATION:
APPLICANT: Takeshi SHIMOMURA et al.
TITLE OF INVENTION: No. 5677164el Protein and Gene Encoding Said Protein
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Penack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch,
MEDIUM TYPE: 500 Kb Storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/448,937A
FILING DATE: May 24, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/148,910
FILING DATE: No. 5677164ember 5, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX: 202-371-8856
TELEX:
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 329 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: human
IMMEDIATE SOURCE:
LIBRARY: Quick-cloneTM human liver cDNA (Clonetechn)
US-08-448-937A-13

Query Match	4.38;	Score 72.2;	DB 1;	Length 329;
Best Local Similarity	59.38;	Pred. No. 5.7e-12;		
Matches 163;	Conservative	0;	Mismatches 103;	Indels 9;
			Gaps	2
QY 1381	TCGCCGACGACTCTCGATGTCGCAAGTCAAGCTGATGGCAACACTTTGTGCACATCCCGC	1440		
Db				
52	TCGAGCTCCCTCGCGGAGGCCCTGGTCCCTCGTCCGCGACCAAGTGCAGCAGCCCT	111		
QY 1441	CACACCTCTMGAGCCACATGATGATTGACACTATGATCTGTGTGAGAGAAATCTTGAGAAACT	1500		
Db				
112	GAGGCTTACGGCGCCGACATCTACGCCCAACATGCTCTGTGCGGCTACTT---GCATGTC	168		
QY 1501	GCGCAAGACACCTCCGACGGGTGACTCTGGAGGCCCCCTGCATCGTGAAGAGCAGCGACC	1560		
Db				
169	AATCTCCGACGCGCTCCGCGGGGACTCAGGGGGGCCCTCGGCTGTGCAGAAAGACGCGTGTG	228		
QY 1561	TACTACGCTCTMGGAATGAGTACTGGGCGCTCGAGATGTGAG-----AAGAGCGACAGG	1614		
Db				
229	GCTTACCTCTTACGGCATCATCAGCTGGGTTAGCGGCTTCGCGCGCTCCACAAAGCCGGGG	288		
QY 1615	GTCCTACCCCAAGTTACCAAAATTCCTGAATTGAT	1649		
Db				
289	GTCCTACCCGCGTGGGCCCAACTGTGGACCTGAT	323		

RESULT 8


```
RESULT 9
US-08-427-640-1
; Sequence 1, Application US/08427640
; Patent No. 5658788
; GENERAL INFORMATION:
; APPLICANT: Berg et al.
; TITLE OF INVENTION: Compounds and Methods for Treatment of Thromboembolic Disor
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/427,640
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/689,410
; FILING DATE: 22 APRIL 1991
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1065 base pairs
; TYPE: Nucleic acid
; STRANDEDNESS: Double
; TOPOLOGY: Linear
; MOLECULE TYPE: DNA
US-08-427-640-1

Query Match          4.3%  Score 72.2; DB 1; Length 1065;
Best Local Similarity 50.5%; Pred. No. 1,1e-11;
Matches 340; Conservative 0; Mismatches 288; Indels 45; Gaps 5;

OY 1028 AGGGCCACTTCTGTGCTGGGGCGCTGATCCACCCCTGCTGGTGTCTCATCTGCGCCACT 1087
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 392 AGCGTTCTCTGTGGGGGCATCTCATCTCTCTCTGATTTCTCTGCGCGCCACT 451
OY 1088 GCACCGA-----CATAAAACCAACATCTTAAAGTGTGTCTAGGGGACCGACCTGA 1141
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 452 GCTTCCAGGAGAGGTTTCCGCCCCACACCTGACGGTGATCTTGGGCACACATACCGGG 511
OY 1142 AGAAGAAGATTTCATGACGAGACCTTTAGGCTGACGAAGATATTCAAGTACAGCCACT 1201
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 512 TGGTCCCTGGCGAGGAGACGAAATTGAAATTCGAAAAATTCATTTGCCATTAAGGAT 571
OY 1202 ACATGAAAGATGATGATTCGCCACAAATGATGATTCCTCAAGTTAAAGCCACTGG 1261
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 572 TCGAT-----GATGACACTTACGACATGTACATTTGGCTGCTGCGACGCTGAATTCGATT 625
OY 1262 ATGTCACGTGTCTAGAAATCAAAATACGTGAAGACTGTGTCTGCTGCTGATGGTCTCT 1321
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 626 CGTCCCGCTGTGCCAGGAGACCGCTGGTCGCGACTGTGTCTCTCCCCGGGGGACC 685
OY 1322 TTCCCT-----CTGGAGTGTAGTCCCAATCTCTGGCTGGGTGTTCACAGAA 1369
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 686 TGCAGCTGCCGAGCTGAGAGAGTGTAGCTCTTCGGCTACGCAAGCATGAGGCGCTTGT 745
OY 1370 CAGGAAAAAGGTCGCCGACCTCTCGATGATGCAAGTCAAGTCAAGTCAAGCTTTGT 1429
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 746 CTCCTTCTATTTCGAGGCGGCTGAAGAGCTCATGTGACAGCTGATACCATTCAGCGCT 805
OY 1430 GCAACTCCCGCCCAACTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1489
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 806 GCACATACAACTTTACTTAAACAGAACAGTCAACCGACATGCTGTGTGTGTGAGACA 865
OY 1490 TTCGAAGAACTGG-----GCAAGACACTTCCAGGAGTACTCTGTGAGGCC 1534
```

```
RESULT 10
US-08-427-640-5
; Sequence 5, Application US/08427640
; Patent No. 5658788
; GENERAL INFORMATION:
; APPLICANT: Berg et al.
; TITLE OF INVENTION: Compounds and Methods for Treatment of Thromboembolic Disor
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/427,640
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/689,410
; FILING DATE: 22 APRIL 1991
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1065 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-427-640-5

Query Match          4.3%  Score 72.2; DB 1; Length 1065;
Best Local Similarity 50.5%; Pred. No. 1,1e-11;
Matches 340; Conservative 0; Mismatches 288; Indels 45; Gaps 5;

OY 1028 AGGGCCACTTCTGTGCTGGGGCGCTGATCCACCCCTGCTGGTGTCTCATCTGCGCCACT 1087
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 392 AGCGTTCTCTGTGGGGGCATCTCATCTCTCTCTGATTTCTCTGCGCGCCACT 451
OY 1088 GCACCGA-----CATAAAACCAACATCTTAAAGTGTGTCTAGGGGACCGACCTGA 1141
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 452 GCTTCCAGGAGAGGTTTCCGCCCCACACCTGACGGTGATCTTGGGCACACATACCGGG 511
OY 1142 AGAAGAAGATTTCATGACGAGACCTTTAGGCTGACGAAGATATTCAAGTACAGCCACT 1201
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 512 TGGTCCCTGGCGAGGAGACGAAATTGAAATTCGAAAAATTCATTTGCCATTAAGGAT 571
OY 1202 ACATGAAAGATGATGATTCGCCACAAATGATGATGATGATGATGATGATGATGATGATG 1261
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 572 TCGAT-----GATGACACTTACGACATGTACATTTGGCTGCTGCGACGCTGAATTCGATT 625
OY 1262 ATGTCACGTGTCTAGAAATCAAAATACGTGAAGACTGTGTCTGCTGCTGATGGTCTCT 1321
```

Db 626 CGTCCCGCTGTGCGCAGAGAGCAGCTGTGTCGACATGTGTGCTTCCCGCGGAGACC 685
QY 1322 TTCCCT-----CTGGAGTGTAGTGCACATCTGTGGCTGGGGTGTATACAGAA 1369
Db 686 TGCACACTCCGAGATGACAGAGTGTGACCTGTCCGCTACGCAAGCATGAGGCTTGT 745
QY 1370 CAGGAAAGGGTCCCGCAGCTCTGTGATGCCAAGTCAAGCATGTATGCCAACACTTGT 1429
Db 746 CTCTCTTTATTTGAGAGGCTCAAGAGGCTCATGTAGACTTACCCATCCAGCCGCT 805
QY 1430 GCAACTCCCGCAACTGTATGACACATGATGATGACAGATGATGCTGCAAGAAATC 1489
Db 806 GCACATCAACATTTACTTAAAGAACAGTACACAGCATCTGTGTGCTGAGACA 865
QY 1490 TTCAGAACCTG-----GCAGACACCTGCGCAGGAGTGTGAGCGCC 1534
Db 866 CTCGAGAGGGGGGCCCGCAGGCAAACTTGCAGAGAGCTGCGCAGGAGATTCGAGAGGCC 925
QY 1535 CCCTGACCTGTGAGAGGAGCGACCTACTACGCTATGAGGATGAGCTGGGCGCTGG 1594
Db 926 CCCTGTGTGTCAAGATGAGCGCGCATGACTTGTGTGGCATCATGAGTGGGCGCTGG 985
QY 1595 AGGTGAGAGAGG-----CCAGGGGTCTACACCCAGTTACCAATTCCTGAATTTGA 1648
Db 986 GCTGTGACAGAGAGATGTCCGGGTGTGTACACCAAGTTACCACTAGACTGGA 1045
QY 1649 TCAAGCCACCAT 1661
Db 1046 TTCGTGACACAT 1058

RESULT 11
US-08-427-640-3
Sequence 3, Application US/08427640
Patent No. 5658788
GENERAL INFORMATION:
APPLICANT: Berg et al.
TITLE OF INVENTION: Compounds and Methods for Treatment of Thromboembolic Disor
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: IN.
COUNTRY: U.S.A.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/427,640
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/669,410
FILING DATE: 22 APRIL 1991
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1068 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-427-640-3

Query Match 4.3%; Score 72.2; DB 1; Length 1068;
Best Local Similarity 50.3%; Pred. No. 1,1e-11;
Matches 340; Conservative 0; Mismatches 288; Indels 45; Gaps 5;
QY 1028 AGGGCCACTTCTGTGTGGGGCGCTGATCCACCCTGCTGGGTGTCTACGTGCTGCCACT 1087

Db 395 AGCGTTCTCTGTGCGGGGSCATTAATCACTCACTGCTGTGATTTCTCTGTCCGCCACT 454
QY 1088 GCACCGA-----CATAAACACAGACATCTAAAGTGTGCTAGGGAGCCAGACCTGA 1141
Db 455 GCTTTCAGAGAGGTTTCCGCGCCACACCTGACGGTGTATCTTGGGCGAAGCATACCGGG 514
QY 1142 AGAAGAGATTTTATTCAGAGAGAGCTTAGGGTGCAGAGATATTCAAGTACAGCCACT 1201
Db 515 TGGTCCCTGCGCAGAGAGAGAGAAATTTGAAGTCCAAAATATCTTCCATAGAGAAAT 574
QY 1202 ACAATGAAAGATGATGATTCGCCAATATGATTTGCTCAAGTTAAAGCCAGTGG 1261
Db 575 TCGAT-----GATGACATTAAGCAATATGATTCGCTGTGCAAGCTGAAACGAT 628
QY 1262 ATGGTACGTGCTGTAGAAATTCGTAAGATACGTGAAGACTGTGTGCTTGGCTGATGGTCT 1321
Db 629 CGTCCGCTGTGCGCAGAGAGAGAGCGGTGCGCACTGTGTGCTTCCCGCGGAGCC 688
QY 1322 TTCCCT-----CTGGAGTGTAGTGCACATCTGTGGCTGGGGTGTATACAGAA 1369
Db 689 TGCAGCTGCGGACTGTGAGAGAGTGTGAGCTTCGCGTACGCGCAAGCATGAGCCCTTGT 748
QY 1370 CAGGAAAGGGTCCCGCAGCTCCTGGATGCCAAAGTCAAGCTGATTTGCCAACCTTGT 1429
Db 749 CTCTCTTTATTTGAGAGCGGCTGAAGAGAGCTCATGTCAAGACTGTACCCATCCAGCCGT 808
QY 1430 GCAACTCCCGCAACTGTATGACACATGATGATGATGACAGTATGATCTGTGCAAGAAATC 1489
Db 809 GCACATCAACATTTACTTAAAGAACAGTACACAGCATCTGTGTGTGTGAGACA 868
QY 1490 TTCAGAACCTG-----GCAGACACCTGCGCAGGAGTGTGAGCGCC 1534
Db 869 CTCGAGAGGGGGGCCCGCAGGCAAACTTGCACAGAGCTGCGCAGGAGCATTCGAGAGGCC 928
QY 1535 CCCTGACCTGTGAGAGGAGCGACCTACTACGCTATGAGATGATGAGCTGGGCGCTGG 1594
Db 929 CCCTGTGTGTGAGAGCATGTGCGCATGACTTGTGTGGCATCATGATGCGGGGCGCTGG 988
QY 1595 AGTGTGAGAGAGG-----CCAGGGGTCTACACCCAGTTACCAATTCCTGAATTTGA 1648
Db 989 GCTGTGACAGAGAGATGTCCGGGTGTGTACACCAAGTTACCACTAGACTGGA 1048
QY 1649 TCAAGCCACCAT 1661
Db 1049 TTCGTGACACAT 1061

RESULT 12
US-09-553-498-9
Sequence 9, Application US/09553498
Patent No. 6309861
GENERAL INFORMATION:
APPLICANT: Ambrosius, Dorothee
APPLICANT: Rudolph, Rainer
APPLICANT: Schaeffner, Joerg
APPLICANT: Schwarz, Elisabeth
TITLE OF INVENTION: Process for the production of naturally folded and secreted pr
FILE REFERENCE: Case 20379
CURRENT APPLICATION NUMBER: US/09/553,498
CURRENT FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: EP99107412.1
PRIOR FILING DATE: 1999-04-26
NUMBER OF SEQ ID NOS: 10
SEQ ID NO 9
LENGTH: 1137
TYPE: DNA
ORGANISM: E. coli
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1137)
US-09-553-498-9

Query Match 4.3%; Score 72.2; DB 4; Length 1137;
Best Local Similarity 50.5%; Pred. No. 1,1e-11;
Matches 340; Conservative 0; Mismatches 288; Indels 45; Gaps 5;

QY 1028 AGGGCCACTTCTGTGTGGGGCCGCTGATCCACCCTGCTGGTCTCATCTGCCCACT 1087
DB 461 AGGGTTCTGTGGGGGSCATCTCATGCTCTGCTGGATCTCTCTCCGCCCACT 520
QY 1088 GCACCGA-----CATAAACACAGACATCTAAAGTGGTGTAGGGACCGACCTGA 1141
DB 521 GCTTCCAGAGAGATTTCCGCCGCCACCTGACGGTGTCTTGGCGAGAACATACCGGG 580
QY 1142 AGAAAGAAATTTATATAGACAGCTTTAGGTGACAGATATTCAAGTACGCCACT 1201
DB 581 TGGTCCCTGGCAGAGAGAGAAATTTGAAGTCAAAAATATCATTCCTCAAGGAAT 640
QY 1202 ACAATGAAGAGATGATTTCCCAATGATTTGATTTGCTCAAGTTAAAGCAGCTG 1261
DB 641 TCGAT-----GATGACACTTACGACATGATTCGGCTGTGAGCTGAAATCGGATT 694
QY 1262 ATGGTCACTGTGCTCTGATCAATCCAAATAGCTGAAGCTGTGCTTGGCTGATGGTCT 1321
DB 695 CGTCCCGCTGTGCTCCAGAGAGACGCTGCTCGGCTGTGCTTCCGCCCGGAGACC 754
QY 1322 TTCCCT-----CTGGAGTGAAGTGCACATCTCTGGCTGGGGTGTACAGAAA 1369
DB 755 TGCAGCTGCGGACTGACGAGATGTGAGCTCTCGGGCTACGGGACGATGAGCCCTGT 814
QY 1370 CAGGAAAGGTCCTCCGCTCTGATGCCAAAGTCAAGCTGATTTGCCAACACTTTGT 1429
DB 815 CTCTCTTCTATTGAGAGGCTGAAGAGCTCATGTGAGCTGATCCATCCAGCCGCT 874
QY 1430 GCACTCCCGCCCACTATGACCATATGATGATGATGATGATGATGATGATGATGATG 1489
DB 875 GCACATCAACACTTTACTTAACAGAACATGATCAGACATGCTGTGTGTGTGTGTGTGT 934
QY 1490 TTCAGAACCTGG-----GCAGACACTGCGCAGGCTGATCTGTGAGGCC 1534
DB 935 CTGAGAGCGGGGCGGCCAGCAAACTTGCAGAGCGCTGCGGAGGGAATTCGGAGGCC 994
QY 1535 CCCTGACTGTGAGAGAGAGCGGACCTACTAGCTATGAGGATAGTGGGCTGG 1594
DB 995 CCCTGCTGTCTGACGATGAGCGGCTGATGCTTGTGTGGCATCATGAGTGGGGCTGG 1054
QY 1595 AGTGTGAGAGAGG-----CCAGGGGTCTACACCCAGTACCAATTTCCTGAATTGGA 1648
DB 1055 GCTGTGACAGAGAGATGCTCCCGGTGTGTACACCAAGGTTACCACTACTAGCTGGA 1114
QY 1649 TCAAGCCACCAT 1661
DB 1115 TTGCTGACACAT 1127

RESULT 13
US-09-618-869-9
; Sequence 9, Application US/0961869
; Patent No. 6453279
; GENERAL INFORMATION:
; APPLICANT: Ambrosius, Dorthée
; APPLICANT: Rudolph, Rainer
; APPLICANT: Schaeffner, Joerg
; APPLICANT: Schwarz, Elisabeth
; TITLE OF INVENTION: PROCESS FOR THE PRODUCTION OF NATURALLY FOLDED AND
; TITLE OF INVENTION: SECRETED PROTEINS BY CO-SECRETION OF MOLECULAR
; TITLE OF INVENTION: CHAPERONES
; FILE REFERENCE: 20381
; CURRENT APPLICATION NUMBER: 05/09/618, 869
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: EP9114811.5
; PRIOR FILING DATE: 1999-07-29
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9

LENGTH: 1137
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1137)
US-09-618-869-9

Query Match 4.3%; Score 72.2; DB 4; Length 1137;
Best Local Similarity 50.5%; Pred. No. 1,1e-11;
Matches 340; Conservative 0; Mismatches 288; Indels 45; Gaps 5;

QY 1028 AGGGCCACTTCTGTGTGGGGCCGCTGATCCACCCTGCTGGTGTCTCATCTGCCCACT 1087
DB 461 AGGGTTCTGTGGGGGSCATCTCATGCTCTGCTGGATCTCTCTCCGCCCACT 520
QY 1088 GCACCGA-----CATAAACACAGACATCTAAAGTGGTGTAGGGACCGACCTGA 1141
DB 521 GCTTCCAGAGAGATTTCCGCCGCCACCTGACGGTGTCTTGGCGAGAACATACCGGG 580
QY 1142 AGAAAGAAATTTATATAGACAGCTTTAGGTGACAGATATTCAAGTACGCCACT 1201
DB 581 TGGTCCCTGGCAGAGAGAGAAATTTGAAGTCAAAAATATCATTCCTCAAGGAAT 640
QY 1202 ACAATGAAGAGATGATTTCCCAATGATTTGATTTGCTCAAGTTAAAGCAGTGG 1261
DB 641 TCGAT-----GATGACACTTACGACATGATTCGGCTGTGAGCTGAAATCGGATT 694
QY 1262 ATGGTCACTGTGCTCTGATCAATCCAAATAGCTGAAGCTGTGCTTGGCTGATGGTCT 1321
DB 695 CGTCCCGCTGTGCTCCAGAGAGACGCTGCTCGGCTGTGCTTCCGCCCGGAGACC 754
QY 1322 TTCCCT-----CTGGAGTGAAGTGCACATCTCTGGCTGGGGTGTACAGAAA 1369
DB 755 TGCAGCTGCGGACTGACGAGATGTGAGCTCTCGGGCTACGGGACGATGAGCCCTGT 814
QY 1370 CAGGAAAGGTCCTCCGCTCTGATGCCAAAGTCAAGCTGATTTGCCAACACTTTGT 1429
DB 815 CTCTCTTCTATTGAGAGGCTGAAGAGCTCATGTGAGCTGATCCATCCAGCCGCT 874
QY 1430 GCACTCCCGCCCACTATGACCATATGATGATGATGATGATGATGATGATGATGATGATG 1489
DB 875 GCACATCAACACTTTACTTAACAGAACATGATCAGACATGCTGTGTGTGTGTGTGTGT 934
QY 1490 TTCAGAACCTGG-----GCAGACACTGCGCAGGCTGATCTGTGAGGCC 1534
DB 935 CTGAGAGCGGGGCGGCCAGCAAACTTGCAGAGCGCTGCGGAGGGAATTCGGAGGCC 994
QY 1535 CCCTGACTGTGAGAGAGAGCGGACCTACTAGCTATGAGGATAGTGGGCTGG 1594
DB 995 CCCTGCTGTCTGACGATGAGCGGCTGATGCTTGTGTGGCATCATGAGTGGGGCTGG 1054
QY 1595 AGTGTGAGAGAGG-----CCAGGGGTCTACACCCAGTACCAAAATTCCTGAATTGGA 1648
DB 1055 GCTGTGACAGAGAGATGCTCCCGGTGTGTACACCAAGGTTACCACTACTAGCTGGA 1114
QY 1649 TCAAGCCACCAT 1661
DB 1115 TTGCTGACACAT 1127

RESULT 14
US-08-811-949-48
; Sequence 48, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIMA, MINEO
; APPLICANT: SATTO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR

```

NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 1314 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1311
US-08-811-949-48

Query Match 4.3%; Score 72.2; DB 2; Length 1314;
Best local Similarity 50.5%; Pred. No. 1.2e-11;
Matches 340; Conservative 0; Mismatches 288; Indels 45; Gaps 5;

QY 1028 AGGGCCACTTCTGTGTGGGCGCTGATCCACCCCTGCTGCTCACTGTCGCCACT 1087
DB 638 AGCGGTCTCTGTGGGGGACATCTACCTCTGCTGATTCCTCTGCGGCCACT 697
QY 1088 GCACCGA-----CATAAAAACAGACATCTAAAGTGTGCTAGGGGACGAGCCTGA 1141
DB 698 GCTTCCAGAGAGATTCCGCCCCACACCTGACGSGTATCTTGGGAGAAACATACCCGG 757
QY 1142 AGAAGAAGATTTCATGAGAGAGCTTTAGGTCAGAAAGATATTCAAGTACAGCCACT 1201
DB 758 TGTGCTCCCTGGCGAGAGAGCAAAATTTGAAGTCGAAAATCATTTGCTATAGGAT 817
QY 1202 ACATGAAAGATGATGATTCCTCCCATGATTTGATGCTGCAAGTAAAGCAGCTGG 1261
DB 818 TCGAT-----GATGACACTTACGACATGATGATTCGCTGCTGACGCTGGAATTCG 871
QY 1262 ATGTCACTGTGCTAGAAATCAATAGCTGAAGACTGTGCTTGCCTGATGGCTCT 1321
DB 872 CGTCCCGCTGTGCCAGAGACACGCTGTCCGCTGCTGCTCTCCCTCCCGCGGAGCC 931
QY 1322 TTCCCT-----CTGGAGTGAAGTGCACATCTCTGAGCTGGGTGTATACAGAAA 1369
DB 932 TGCAGCTGCCGAGCTGAGAGCTGTGAGCTCTCCGGCTACGCGAAGCATGAGCGCTTGT 991
QY 1370 CAGAAAAAGGTCGCCAGCTCTGATGCGCAAGTCAAGTGAATGCAACACTTTGT 1429
DB 992 CTCCTTTCTATTGAGCGCGCTGAGAGGCTCATGTCAACACTGTCACCATCCACCCCT 1051
QY 1430 GCAACTCCCGGCAAGCTCTAAGCAACATGATTTGAGTGAAGTATGATCTGTGAGCAATC 1489
DB 1052 GCACATCAACAATTACTTAACAGAGAGTACCGGCAACATGCTGTGTGTGAGACA 1111

```

```

QY 1490 TTCAAAACCTGC-----GCAAGACACTCCAGGCTGACTCTGAGAGCC 1534
DB 1112 CTCGAGAGGCGGGGCCCGAGGAACTTGACAGAGCTCCAGGGGCTTGAGGAGCC 1171
QY 1535 CCGTACCTGTGAGAGGAGGACGACCTACTAGCTGTATGGATAGTGAAGTGGGCGCTG 1594
DB 1172 CCTGTGTGTGTGAGAGGATGGCCGATCACTTTGTGTGGCATCATCACTGCGGCTG 1231
QY 1595 AGTGTGAGAGAG-----CAGAGGCTCTACACCACTTACCAATTCCTGAATTGA 1648
DB 1232 GCTGTGAGCAAGAGATGTCGCGGTGTGTACAAAGTTACCAACTACAGACTGA 1291
QY 1649 TCAAAGCCACCAT 1661
DB 1292 TTCTGACACAT 1304

RESULT 15
US-08-883-795A-39
Sequence 39, Application US/08883795A
Patent No. 5985607
GENERAL INFORMATION:
APPLICANT: Delcuve, Genevieve
TITLE OF INVENTION: Recombinant DNA Molecules and Expression
TITLE OF INVENTION: Vectors for Tissue Plasmogen Activator
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: BERESKIN & PARR
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/883,795A
FILING DATE: 27-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gravelle, Micheline
REGISTRATION NUMBER: 40,261
REFERENCE/DOCKET NUMBER: 7841-062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 1955 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-883-795A-39

Query Match 4.3%; Score 72.2; DB 2; Length 1955;
Best local Similarity 50.5%; Pred. No. 1.5e-11;
Matches 340; Conservative 0; Mismatches 288; Indels 45; Gaps 5;

QY 1028 AGGGCCACTTCTGTGTGGGCGCTGATCCACCCCTGCTGCTCACTGTCGCCACT 1087
DB 1103 AGCGGTCTCTGTGGGGGAGATCTCATGACTCTGCTGATTTCTCTGCGGCCACT 1162
QY 1088 GCACCGA-----CATAAAAACAGACATCTAAAGTGTGCTAGGGGACGAGACTGA 1141
DB 1163 GCTTCCAGAGAGGTTTCCGCCCCACACCTGACGSGTATCTTGGGAGAAATACCGGG 1222

```


GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 18:44:22 ; Search time 127.5 Seconds

(without alignments)
8816.543 Million cell updates/sec

Title: US-09-912-559-2

Perfect score: 1683
Sequence: 1 atgtttccagcagatgctcga.....aaagtgaaagtgctcttaa 1683

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 478924 seqs, 333959956 residues

Total number of hits satisfying chosen parameters: 957848

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*\n2: /cgn2_6/ptodata/2/pubpna/PTCT_NEW_PUB.seq:*\n3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*\n4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*\n5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*\n6: /cgn2_6/ptodata/2/pubpna/PTCTUS_PUBCOMB.seq:*\n7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*\n8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*\n9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*\n10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*\n11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*\n12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*\n13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*\n14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1683	100.0	1683	10	US-09-912-559-2
2	1679.8	99.8	1683	10	US-09-912-559-1
3	1679.8	99.8	3008	10	US-09-880-107-1668
C 4	120.2	7.1	451	10	US-09-864-761-11164
C 5	117	7.0	117	10	US-09-864-761-27791
6	114.4	6.8	2036	10	US-09-954-456-552
7	114.4	6.8	2036	10	US-09-880-107-1612
8	72.8	4.3	614	10	US-09-879-792-33
9	72.2	4.3	1689	10	US-09-969-271-6
10	72.2	4.3	2519	10	US-09-969-271-5
11	72.2	4.3	2641	9	US-09-974-298-144
12	66.4	3.9	1614	10	US-09-888-615-45
13	66.4	3.9	1748	10	US-09-879-792-11
14	66.4	3.9	2412	9	US-09-978-295A-68
15	66.4	3.9	2412	9	US-09-978-697-68
16	66.4	3.9	2412	9	US-09-978-192A-68
17	66.4	3.9	2412	9	US-09-999-832A-68
18	66.4	3.9	2412	9	US-09-978-189-68
19	66.4	3.9	2412	9	US-10-174-590-63

20	66.4	3.9	2412	9	US-10-176-758-63	Sequence 63, Appl
21	66.4	3.9	2412	9	US-10-175-737-63	Sequence 63, Appl
22	66.4	3.9	2412	9	US-10-173-706-63	Sequence 63, Appl
23	66.4	3.9	2412	9	US-10-175-706-63	Sequence 63, Appl
24	66.4	3.9	2412	9	US-10-175-752-63	Sequence 63, Appl
25	66.4	3.9	2412	9	US-10-176-482-63	Sequence 63, Appl
26	66.4	3.9	2412	9	US-10-176-752-63	Sequence 63, Appl
27	66.4	3.9	2412	9	US-10-176-913-63	Sequence 63, Appl
28	66.4	3.9	2412	9	US-10-180-552-63	Sequence 63, Appl
29	66.4	3.9	2412	9	US-10-180-557-63	Sequence 63, Appl
30	66.4	3.9	2412	9	US-10-173-700-63	Sequence 63, Appl
31	66.4	3.9	2412	9	US-10-174-572-63	Sequence 63, Appl
32	66.4	3.9	2412	9	US-10-174-572-63	Sequence 63, Appl
33	66.4	3.9	2412	9	US-10-174-582-63	Sequence 63, Appl
34	66.4	3.9	2412	9	US-10-174-588-63	Sequence 63, Appl
35	66.4	3.9	2412	9	US-10-175-738-63	Sequence 63, Appl
36	66.4	3.9	2412	9	US-10-175-740-63	Sequence 63, Appl
37	66.4	3.9	2412	9	US-10-175-743-63	Sequence 63, Appl
38	66.4	3.9	2412	9	US-10-176-488-63	Sequence 63, Appl
39	66.4	3.9	2412	9	US-10-176-492-63	Sequence 63, Appl
40	66.4	3.9	2412	9	US-10-176-747-63	Sequence 63, Appl
41	66.4	3.9	2412	9	US-10-176-750-63	Sequence 63, Appl
42	66.4	3.9	2412	9	US-10-176-985-63	Sequence 63, Appl
43	66.4	3.9	2412	9	US-10-176-987-63	Sequence 63, Appl
44	66.4	3.9	2412	9	US-10-176-991-63	Sequence 63, Appl
45	66.4	3.9	2412	9	US-10-176-992-63	Sequence 63, Appl

ALIGNMENTS

RESULT 1
US-09-912-559-2
; Sequence 2, Application US/09912559
; Patent No. US20020142316A1
; GENERAL INFORMATION:
; APPLICANT: ROEMTCH, JUERGEN
; APPLICANT: STOEHR, HANS-ARNOLD
; APPLICANT: FEUSSNER, ANNETTE
; APPLICANT: LANG, WIEGAND
; APPLICANT: WEIMER, THOMAS
; APPLICANT: BECKER, MARGRET
; APPLICANT: NERLICH, CLAUDIA
; APPLICANT: MUTH-NAUMANN, GUDRUN
; TITLE OF INVENTION: MUTANTS OF THE FACTOR VII-ACTIVATING PROTEASE AND
; TITLE OF INVENTION: DETECTION METHODS USING SPECIFIC ANTIBODIES
; FILE REFERENCE: 06478.1457
; CURRENT APPLICATION NUMBER: US/09/912.559
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: DE 100 36 641.4
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: DE 100 50 040.4
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: DE 100 52 319.6
; PRIOR FILING DATE: 2000-10-21
; PRIOR APPLICATION NUMBER: DE 101 18 706.8
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1683
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-912-559-2

Query Match 100.0%; Score 1683; DB 10; Length 1683;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1683; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTTTCCAGCAGATGCTGATCTTCATGTTCTGCTGTTAATGAGCTCTGGTGGAAAGACA 60
DB 1 ATGTTTCCAGCAGATGCTGATCTTCATGTTCTGCTGTTAATGAGCTCTGGTGGAAAGACA 60

QY	61	GCCTGGGGGTTCCCTCATCTCTTTTATTGGAAAGCTGTGAGCCAGACTGGACCCCTGAC	120
Db	61	GCCCTGGGGGTTCCCTCATCTCTTTTATTGGAAAGCTGTGAGCCAGACTGGACCCCTGAC	120
QY	121	CAGTATGATTACAGCTACGAGAGATTATTAATCAGAGAGAGAACACAGACAGTACACTTAC	180
Db	121	CAGTATGATTACAGCTACGAGAGATTATTAATCAGAGAGAGAACACAGACAGTACACTTAC	180
QY	181	CATGCTGAGAAATCCGACTGGTACTATACAGAGAGAACAAAGCTGATCCATGCGACGCCAC	240
Db	181	CATGCTGAGAAATCCGACTGGTACTATACAGAGAGAACAAAGCTGATCCATGCGACGCCAC	240
QY	241	CCCTGTGAACACAGGTGGGGAGCTGCCCTGTCATAGGAGAGACCTTCACATCCACTGCTG	300
Db	241	CCCTGTGAACACAGGTGGGGAGCTGCCCTGTCATAGGAGAGACCTTCACATCCACTGCTG	300
QY	301	GCCTCTTTCCTGSGGAATAATGTCACAAAGTSCAAATATAGTGCNAAAGCAACCCATGT	360
Db	301	GCCTCTTTCCTGSGGAATAATGTCACAAAGTSCAAATATAGTGCNAAAGCAACCCATGT	360
QY	361	GGCCGGGGCCAAATGTCCTATTACCCAGAGTCCCTCCCTACCGCGTGTGTGTAAACAC	420
Db	361	GGCCGGGGCCAAATGTCCTATTACCCAGAGTCCCTCCCTACCGCGTGTGTGTAAACAC	420
QY	421	CCCTTACACAGTCCCAAGTGTCTCCCAAGTGTGTTCCGTATGAGAGCCAAACCCCTGCGAG	480
Db	421	CCCTTACACAGTCCCAAGTGTCTCCCAAGTGTGTTCCGTATGAGAGCCAAACCCCTGCGAG	480
QY	481	AATGGGGCTAAGTCTGCTCCCGCATAGCGAGATCCAAAGTTCACTGCTGTCTCCGAC	540
Db	481	AATGGGGCTAAGTCTGCTCCCGCATAGCGAGATCCAAAGTTCACTGCTGTCTCCGAC	540
QY	541	CAGTTAAAGGGGAAATTTGTGAATATAGTTCGATGATCTGCTATGTTGGCATAGCTCAC	600
Db	541	CAGTTAAAGGGGAAATTTGTGAATATAGTTCGATGATCTGCTATGTTGGCATAGCTCAC	600
QY	601	TCTTACCAGGGGAAATATATAGACAGTCAACCAAGATCGGTGCTTTACTGGAATCTC	660
Db	601	TCTTACCAGGGGAAATATATAGACAGTCAACCAAGATCGGTGCTTTACTGGAATCTC	660
QY	661	CACCTCCTCTTCCAGAGAAATATACATATGTTTATGAGAGATGCTGAACCCATGGAGATT	720
Db	661	CACCTCCTCTTCCAGAGAAATATACATATGTTTATGAGAGATGCTGAACCCATGGAGATT	720
QY	721	GGGGAAACAATTTGTGAGAAACCCAGATGGGAGCAAAAGCCCTGGTCTTATTAA	780
Db	721	GGGGAAACAATTTGTGAGAAACCCAGATGGGAGCAAAAGCCCTGGTCTTATTAA	780
QY	781	GTTTACCAATGACAAAGGTGAAAATGGGAATACTGTGATGTCTCAGCCTCTCAGCCAGAC	840
Db	781	GTTTACCAATGACAAAGGTGAAAATGGGAATACTGTGATGTCTCAGCCTCTCAGCCAGAC	840
QY	841	GTTTGCTTACCAGAGAAAGCCCACTGAGCCCATCAACCAAGCTCCGGGGTTTGACTCC	900
Db	841	GTTTGCTTACCAGAGAAAGCCCACTGAGCCCATCAACCAAGCTCCGGGGTTTGACTCC	900
QY	901	TGTGGAAGACTGAGATAGCAGAGAGAAAGATCAAGAGAATCTATGAGGCTTTAAGAC	960
Db	901	TGTGGAAGACTGAGATAGCAGAGAGAAAGATCAAGAGAATCTATGAGGCTTTAAGAC	960
QY	961	ACGGCGGGCAACCAACCTGGCAGGGGCTCCCTCAAGTCCCTGGCTGTGACAACTTCC	1020
Db	961	ACGGCGGGCAACCAACCTGGCAGGGGCTCCCTCAAGTCCCTGGCTGTGACAACTTCC	1020
QY	1021	ATGCCCCAGGGCCACTTCTGTGTGTGGGGCGCTGATCAACCCCTCTGTGGGTCACTGCT	1080
Db	1021	ATGCCCCAGGGCCACTTCTGTGTGTGGGGCGCTGATCAACCCCTCTGTGGGTCACTGCT	1080
QY	1081	GGCCACTGACGACATPAAAAACCAAGCACTTAAGAGTGTGTGTAAGGGGACAGAGACTG	1140
Db	1081	GGCCACTGACGACATPAAAAACCAAGCACTTAAGAGTGTGTGTAAGGGGACAGAGACTG	1140
QY	1141	AAGAAAGAGAAATTCATGACAGAGACTTTAGGTGTCAGAAAGATATTCAGATACAGCCAC	1200

Db	1141	AAAGAAAGAAATTTTCATGACACAGCTTTAGGCTGCAGAGAAATATTCAGTACACCCAC	1200
Qy	1201	TACAAAGAAAGAGATGAGATATCCCCACAATGATATGCAATTGCTCAAGTTAAAGCCAGTG	1260
Db	1201	TACAAAGAAAGAGATGAGATATCCCCACAATGATATGCAATTGCTCAAGTTAAAGCCAGTG	1260
Qy	1261	GATGTCACCTGTCGTCTGCAATCCAAATACGTGAGACTGTGTGCTTGCCTGATGGGTCC	1320
Db	1261	GATGTCACCTGTCGTCTGCAATCCAAATACGTGAGACTGTGTGCTTGCCTGATGGGTCC	1320
Qy	1321	TTTCCCTCTGGAGAGTGGCACACTCTGCTGGCTGGGGTGTATCAGAAACAGAAAGAG	1380
Db	1321	TTTCCCTCTGGAGAGTGGCACACTCTGCTGGCTGGGGTGTATCAGAAACAGAAAGAG	1380
Qy	1381	TCCCGCAGCTCTGTGGATGGCCAAAGTCACAGTGAATGGCCAACTTTTGTGCATCCGGC	1440
Db	1381	TCCCGCAGCTCTGTGGATGGCCAAAGTCACAGTGAATGGCCAACTTTTGTGCATCCGGC	1440
Qy	1441	CAACTATGACACACATGATTTGATGACAGATGATCTGTGCAAGAAATCTTCAGAAACT	1500
Db	1441	CAACTATGACACACATGATTTGATGACAGATGATCTGTGCAAGAAATCTTCAGAAACT	1500
Qy	1501	GGGGAAGACACCTGGCAGAGGTGATCTGTGAGAGCCCGCTGACCTGTGAGAGAGCCGACAC	1560
Db	1501	GGGGAAGACACCTGGCAGAGGTGATCTGTGAGAGCCCGCTGACCTGTGAGAGAGCCGACAC	1560
Qy	1561	TACTAGCTATGAGATAGTAGAGCTGGGGCTGTGAGTGTGAGAAAGAGCCAGGGGTAC	1620
Db	1561	TACTAGCTATGAGATAGTAGAGCTGGGGCTGTGAGTGTGAGAAAGAGCCAGGGGTAC	1620
Qy	1621	ACCAAGTTACCAATTTCTGAATTTGGATCAAAAGCCACCATCAAAATGAAATGGCTTC	1680
Db	1621	ACCAAGTTACCAATTTCTGAATTTGGATCAAAAGCCACCATCAAAATGAAATGGCTTC	1680
Qy	1681	TAA 1683	
Db	1681	TAA 1683	
RESULT 2			
US-09-912-559-1			
Sequence 1. Application US/09912559			
Patent No. US20020142316A1			
GENERAL INFORMATION:			
APPLICANT: KOEMISCH, JUERGEN			
APPLICANT: STOEHR, HANS-ARNOOLD			
APPLICANT: FEUSSNER, ANNETTE			
APPLICANT: LANG, WIEGAND			
APPLICANT: WEIMER, THOMAS			
APPLICANT: BECKER, MARCERET			
APPLICANT: NERLICH, CLAUDIA			
APPLICANT: MOTH-NHOMANN, GODRUN			
TITLE OF INVENTION: MUTANTS OF THE FACTOR VII-ACTIVATING PROTEASE AND			
FILE REFERENCE: 06478.1457			
CURRENT APPLICATION NUMBER: US/09/912.559			
CURRENT FILING DATE: 2001-07-26			
PRIOR APPLICATION NUMBER: DE 100 36 641.4			
PRIOR FILING DATE: 2000-07-26			
PRIOR APPLICATION NUMBER: DE 100 50 040.4			
PRIOR FILING DATE: 2000-10-10			
PRIOR APPLICATION NUMBER: DE 100 52 319.6			
PRIOR FILING DATE: 2000-10-21			
PRIOR APPLICATION NUMBER: DE 101 18 706.8			
PRIOR FILING DATE: 2001-04-12			
NUMBER OF SEQ ID NOS: 4			
SOFTWARE: PatentIn Ver. 2.1			
SEQ ID NO 1			
LENGTH: 1683			
TYPE: DNA			
ORGANISM: Homo sapiens			
US-09-912-559-1			

Query Match	99.8%	Score 1679.8;	DB 10;	Length 1683;
Best Local Similarity	99.9%	Pred. No. 0;		
Matches 1681; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

QY	1	ATGTTTCCAGAGATGTCTGATCTCCATGTTCTCTGTTTATGGCTCTGTGGGAAGACA	60
Db	1	ATGTTTCCAGAGATGTCTGATCTCCATGTTCTCTGTTTATGGCTCTGTGGGAAGACA	60
QY	61	GCCGTGGGTTCCCTCATGTCTTTTATGGAAAGCCTGGACCCGAGCTGGACCCCTGCAC	120
Db	61	GCCGTGGGTTCCCTCATGTCTTTTATGGAAAGCCTGGACCCGAGCTGGACCCCTGCAC	120
QY	121	CAGTATGATTAAAGCTACGAGAGATTATATTCAGAAAGACACCAGTACACTTACC	180
Db	121	CAGTATGATTAAAGCTACGAGAGATTATATTCAGAAAGACACCAGTACACTTACC	180
QY	181	CATGCTGAGATTCCTGACTGGTACTTACACTGAGAGACCAAGCTGATCCATGGACCCCAAC	240
Db	181	CATGCTGAGATTCCTGACTGGTACTTACACTGAGAGACCAAGCTGATCCATGGACCCCAAC	240
QY	241	CCCTGTGAACACGGTGGGAGTGCCTGTCTCATGGAGCACCCTACATGACGTGCTG	300
Db	241	CCCTGTGAACACGGTGGGAGTGCCTGTCTCATGGAGCACCCTACATGACGTGCTG	300
QY	301	GCCTCTTTCTGGGAATTAAGTGTCAAGAAAGTCAAAATACCTGGAAGGACAACCCATGT	360
Db	301	GCCTCTTTCTGGGAATTAAGTGTCAAGAAAGTCAAAATACCTGGAAGGACAACCCATGT	360
QY	361	GGCCGGGGCCCATGTCTCATTAACCCAGAGTCTCCCTACTACCGCTGTGTGTGAACAC	420
Db	361	GGCCGGGGCCCATGTCTCATTAACCCAGAGTCTCCCTACTACCGCTGTGTGTGAACAC	420
QY	421	CCCTTACACAGATCCCAAGCTGTCCCAAGTGGTCTGTATGGAGGCCAACCCTGCGAC	480
Db	421	CCCTTACACAGATCCCAAGCTGTCCCAAGTGGTCTGTATGGAGGCCAACCCTGCGAC	480
QY	481	AATGGGGCTACTCTGCTCCCGCATTAAGCGAGATCCAAAGTTACCTTGCTGCTCCGCAC	540
Db	481	AATGGGGCTACTCTGCTCCCGCATTAAGCGAGATCCAAAGTTACCTTGCTGCTCCGCAC	540
QY	541	CAGTTCAAGGGGAATTTCTGGAATAGTGTGATGACTGCTATGTTGGCGATGGCTAC	600
Db	541	CAGTTCAAGGGGAATTTCTGGAATAGTGTGATGACTGCTATGTTGGCGATGGCTAC	600
QY	601	TCTTACCGAGGGAATTAATAGACAGTCAACACAGATGCTGTGCTTTTACTGGAATCC	660
Db	601	TCTTACCGAGGGAATTAATAGACAGTCAACACAGATGCTGTGCTTTTACTGGAATCC	660
QY	661	CACCTCTTTTCAGAGAGAAATTAACACATGTTTATAGGAGATGCTGGAACCCATGGGAAT	720
Db	661	CACCTCTTTTCAGAGAGAAATTAACACATGTTTATAGGAGATGCTGGAACCCATGGGAAT	720
QY	721	GGGGAAACCAATTTCTGCGAANAACCAAGATCGGACGAGAAAGCCCTGTGCTTTATTA	780
Db	721	GGGGAAACCAATTTCTGCGAANAACCAAGATCGGACGAGAAAGCCCTGTGCTTTATTA	780
QY	781	GTTACCAATGACAGAGTGAATGGGAATACTGTGATGTCTCAGCTGTCTAGGCCAGGAC	840
Db	781	GTTACCAATGACAGAGTGAATGGGAATACTGTGATGTCTCAGCTGTCTAGGCCAGGAC	840
QY	841	GTTTGCTTACCCAGAGAAAGCCCACTGAGCCATTAACCAAGCTTCCGGGGTTTGACTCC	900
Db	841	GTTTGCTTACCCAGAGAAAGCCCACTGAGCCATTAACCAAGCTTCCGGGGTTTGACTCC	900
QY	901	TGTGGAAGACGATAGAGAGAGAGAAAGATCAAGACAATCTTGGAGGCTTTAAGAGC	960
Db	901	TGTGGAAGACGATAGAGAGAGAGAAAGATCAAGACAATCTTGGAGGCTTTAAGAGC	960
QY	961	ACGGGGGGAAGCACCATGGCAGGGGTCCCTCTCAGTCTTGGCTGTGACCATTTCC	1020
Db	961	ACGGGGGGAAGCACCATGGCAGGGGTCCCTCTCAGTCTTGGCTGTGACCATTTCC	1020

OY	1021	ATGCGCCAGGGCCACCTCGTGGTGGGGGGCGGTATCAGCACCCCTGCTGGGTGCTCACTGCT	1080
Db	1021	ATGCGCCAGGGCCACTCTCTGTGGTGGGGGGCGGTATCAGCACCCCTGCTGGGTGCTCACTGCT	1080
OY	1081	GCCCACTGCACCGACATATAAAACACAGACATCTAAAGGTGTGTCTAGGGGGACACGACCTG	1140
Db	1081	GCCCACTGCACCGACATATAAAACACAGACATCTAAAGGTGTGTCTAGGGGGACACGACCTG	1140
OY	1141	AAGAAAGAGAAATTTTCATGACACAGCTTTAGGTGGCAGAAAGATATTCAAGTACAGCCAC	1200
Db	1141	AAGAAAGAGAAATTTTCATGACACAGCTTTAGGTGGCAGAAAGATATTCAAGTACAGCCAC	1200
OY	1201	TACAAATGAAAGAGATGAGATTCCCCACCAATGATATTGCATTGCTCAAGTTAAAGCAGTG	1260
Db	1201	TACAAATGAAAGAGATGAGATTCCCCACCAATGATATTGCATTGCTCAAGTTAAAGCAGTG	1260
OY	1261	GATGTCACCTGTGCTCTGTGAATTCCAATACGTGAAGACGTGTGCTGCTGATGTGGCTC	1320
Db	1261	GATGTCACCTGTGCTCTGTGAATTCCAATACGTGAAGACGTGTGCTGCTGATGTGGCTC	1320
OY	1321	TTTTCCTCTGTGGAGTGAATGCACATCTCTGTGCTGGGGTGTTCAGAAACAGAGAAAGG	1380
Db	1321	TTTTCCTCTGTGGAGTGAATGCACATCTCTGTGCTGGGGTGTTCAGAAACAGAGAAAGG	1380
OY	1381	TCCCGCAGCTCTGTGGATGGCCAAAGTCAGCTGATTTGGCACACTTTGTGCCAATCTCCGC	1440
Db	1381	TCCCGCAGCTCTGTGGATGGCCAAAGTCAGCTGATTTGGCACACTTTGTGCCAATCTCCGC	1440
OY	1441	CAACTCTATGACCAATGATGATGACAGTATGATCTGTGCAGGAAATCTTCAGAAACT	1500
Db	1441	CAACTCTATGACCAATGATGATGACAGTATGATCTGTGCAGGAAATCTTCAGAAACT	1500
OY	1501	GGGCAAGACACCTGTGCAGGGTGAATCTGGAGGGCCCCCTGACCTGTGGAAAGAGCGGCAAC	1560
Db	1501	GGGCAAGACACCTGTGCAGGGTGAATCTGGAGGGCCCCCTGACCTGTGGAAAGAGCGGCAAC	1560
OY	1561	TACTACGCTTATGGATATGATGAGCTGGGGCTGTGGAGTGTGAAGAGACCCAGAGGGCTTAC	1620
Db	1561	TACTACGCTTATGGATATGATGAGCTGGGGCTGTGGAGTGTGAAGAGACCCAGAGGGCTTAC	1620
OY	1621	ACCAAGATTACCAATTTCTCTAATTGGATCAAAAGCCACATCAAAAGTGAAAGTGGCTTC	1680
Db	1621	ACCAAGATTACCAATTTCTCTAANTTGGATCAAAAGCCACATCAAAAGTGAAAGTGGCTTC	1680
OY	1681	TAA 1683	
Db	1681	TAA 1683	
RESULT 3			
US-09-880-107-1668			
Sequence 1668: Application US/09880107			
Patient No. US20020142961A1			
GENERAL INFORMATION:			
APPLICANT: Horne, Darci T.			
APPLICANT: Vockley, Joseph G.			
APPLICANT: Scherf, Uwe			
APPLICANT: Gene Logic, Inc.			
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer			
FILE REFERENCE: 44921-5028-WO			
CURRENT APPLICATION NUMBER: US/09/880,107			
CURRENT FILING DATE: 2001-06-14			
PRIOR APPLICATION NUMBER: US 60/211,379			
PRIOR FILING DATE: 2000-06-14			
PRIOR APPLICATION NUMBER: US 60/237,054			
PRIOR FILING DATE: 2000-10-02			
NUMBER OF SEQ ID NOS: 3950			
SOFTWARE: PatentIn Ver. 2.1			
SEQ ID NO 1668			
LENGTH: 3008			
TYPE: DNA			
ORGANISM: Homo sapiens			
FEATURE:			


```
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 683290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 552
; LENGTH: 2036
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-552

Query Match      6.8%; Score 114.4; DB 10; Length 2036;
Best Local Similarity 49.6%; Pred. No. 7.2e-25;
Matches 605; Conservative 0; Mismatches 561; Indels 54; Gaps 10;
```

```
Db 1213 CTGGCCGACAGTTCATTCGCGGGGCTCTCCGCTGCGCGGCGTGCACCCCTG----- 1265
Qy 1006 CCTCTGACCATTCCTCCATTCGCCCGCCACTTCTGTGTGGGGCGCTTCACCCCTGC 1065
Db 1266 --GCTGGCCGCCATTCATTCGCGGACAGCTTCTGGCGCGGAGCGCTGTCCACACCTGC 1223
Qy 1066 TGGGTGTCACCTGCTGCCCGACCTGCACCTAATAAACGACATCTAA-----AGGTG 1119
Db 1334 TGGGTGCTGTGCGCCCGCCCGACCTGCTTCCACAGCCCGCCCGACAGCGCTCTCGCTG 1383
Qy 1120 GTGCTAAGGGGACCGACCTGTAAGAAAGAAATTTCTATGAGCAGCTTTAGGGTCAG 1179
Db 1384 GTGCTGCGGCGACGACTTCTTCAACCGCACGAGCGAGCTGACGAGCTTCGGGATGAG 1443
Qy 1180 AAGTATTTCAAGTACAGCGCCACTACATGAAGAAAGATGAGATTTCCCATATGATTTGCA 1239
Db 1444 AAGTACATCCCGTACACCTGTACTCGGTTCACCCCGAG--CGACCACGACCTTGTC 1500
Qy 1240 TTGCTCAAGTTAAAGCCAGTGAATGTCATCTGCTCTAGAAATCCAAATACGTGAACAT 1299
Db 1501 CTGATCCGCGCTGAAGAAAGAGGAGCGCTGTGCCACAGCGCTCGACATTCGTGACGCC 1560
Qy 1300 GTGCTGCTGGCTGATG-----GGTCTTCCCTCTGGGAGTGAATGTCACATCTGCGC 1353
Db 1561 ATCTGCTGCGCGAGCGCGGACGACACCTTCCCGCAGAGACACAATGSCAATTTGCGGGC 1620
Qy 1354 TGGGG-----TGTTACAGAAACAGGAAAGGTTCCCGCAGCTCTCTGATGCCAAGATC 1407
Db 1621 TGGGGCCACTTGGATGAGAACGTGAGCGGCTACTCCAGCTCCCTCGGAGGCGCTGTC 1680
Qy 1408 AAGCTGATTTCCACACCTTTGTGCAACCTCCGCCACCTCTATGACCATGATGATGATGAC 1467
Db 1681 CCCCTGGTCCCGACACACAAAGTGAAGAGCGCTTGAAGCTTACAGCGCTCGACATAGCCCC 1740
Qy 1468 AGTATGCTGTGTCAGGAAATCTTCAGAAACTGGGGCAAGACCTTGCAGGAGTACTT 1527
Db 1741 AACATCTGTGTGCGCGGCTACTT---CGACTGCACAGTCCGACCGCTGCAGCGGAGCTCA 1797
Qy 1528 GGAAGGCCCTCTACCTGTGAGAGAGGACCACTTACTACGTATAGGATATGATGATG 1587
Db 1798 GGGGGGCGCCCTGGCTGTGAGAGAAAGCGGTGAGCTTACTTACGGCATATCAGCTGG 1857
Qy 1588 GCGCTGAGTGTGAG-----AAGAGCCAGGGGTCTACACCCAGTTACCAATTCCTG 1641
Db 1858 GGTGAGCGCTGGGGCGGCTCCACACAGCCGGGGGTCTACACCCGCTGGCCAACTATGTG 1917
Qy 1642 AATTGATCAAAAGCCCAT 1661
Db 1918 GACTGATCAAGACCGAT 1937

RESULT 7
US-09-880-107-1612
; Sequence 1612, Application US/09880107
; Patent No. US2002012981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 1612
; LENGTH: 2036
```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 D14012
US-09-880-107-1612

```

```

Query Match      6.8%; Score 114.4; DB 10; Length 2036;
Best Local Similarity 49.6%; Pred. No. 7.2e-25;
Matches 605; Conservative 0; Mismatches 561; Indels 54; Gaps 10;

```

```

Oy 469 ACCCCGTCAGATAGGAGGCTACCTGCTCCCGGCAATACGAGATCCAACTTCACTGT 528
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 745 AGCCCTTGGCTTACAGCGGAGCACTGCCACCTGATGTGGCCACCGGACCACTGTGT 804
Oy 529 GCGCTGCCAGCAGTTCAGAGGAAATTCGTGAATAGTTCTGATGA---CTGCTAT 585
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 805 GCGTCCCAACAGAGCTTCGTGAGCGGCTCTGCACATCGACCTGATGACGCTGCTC 864
Oy 586 GTTGGGAGATGGCACTTACCGSAGGAAATGATGAGCATCAACACGACATGCTGC 645
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 865 TTGGGGAGAGCGCATGCGGTACCGTGGCGTGGCCAGCACTCAAGCTCGGCTGAGCTGC 924
Oy 646 CTTTACTGACATCCCACTCTCTTSCAGAGAGATTTACACATGTTTATGAGATGCT 705
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 925 CTGGCTTGGAGCTCGGATCTGCTTACAGAGAGCTCAAGTGAATCCGTCGGCGCGC 984
Oy 706 GAAACCATGGGATTTGGGAGACACATTTCTGCAAGAACCCGATGGGAGGAAAGCC 765
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 985 GCGCTGCTGGGCTGGGCGCCCATGCTTACTGCCGAAATCCGGAATGAGAGAGGCCC 1044
Oy 766 TGTGCTTTATTAAGATTACCAATGACAAAGTGAATGGAAATGATGTCTGACGC 825
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1045 TGTGTGCTAGCT---GCTAAGAGACAGCGCTCTGCGGAAATGATGCGCGCTGAGAGCC 1101
Oy 826 TGTCAAGCCAGAGAGGTTGCTTACCCAGAGAAAGCCCACTGACCATCAACCACTT 885
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1102 TGGGATTCCTT-----CACCAAGATCCACATCTCACCGGATCTCTGCGACCTTG 1152
Oy 886 CCGGGGTTTACTCTCTGTGAAAGACTGAGATAGCAGAGAGAGATCAAGAAATCTAT 945
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1153 CCGAGCCAGCGCTCCCGGGGCGCAGGCTCGGCGAGAGGACACAAAGAGAGAGCTTG 1212
Oy 946 GGAGGCTTTAAGAGCAGCGGCGGAGACACCATGAGGAGGCTCTCATCTCTGCTG 1005
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1213 CTGGGGCCAGATATATGAGGGGCTCTCTCTCTGCTCGCGCTCGCACCCCTG----- 1265
Oy 1006 CTTCTGACATCTCCATGCCCCAGGGCCACTTCTGTGTGGGCGCTGATCAACCCCTGC 1065
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1266 --GCTGGCGGCATCTACATCGGGGAGCAGCTTCTGCGCGGAGGCTGTGTCACACCTGC 1323
Oy 1066 TGGGTGCTACTGTGCCACTTGCACGACATTAANAACGACATCTTA-----AGGTG 1119
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1324 TGGGTGTGTGTCGGCGCCGACCTTCTTCCACAGCCCGCCGAGGAGACGCTTCCGTG 1383
Oy 1120 GTCTAGGGGACCAAGACTTGAAGAAAGAAATTTTCATGAGCAGAGCTTTAGGGTGCA 1179
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1384 GTCTGGGCGACACTTCTTCACCGCAGCAGAGAGAGTACGACATCTTGAGATGAG 1443
Oy 1180 AAGATTTCAATACAGCCACTTACATGAAGAGATGAGATTTCCCAATGATATGCA 1239
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1444 AATATATCCCTTACACCTGTACTGTGTTTACACCCAG---CGACCAAGACCTTGTC 1500
Oy 1240 TTCTCTAAGTTAAAGCAAGTATGATGTGCTCTTGAATTCCAATATGAGAAAGCT 1299
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1501 CTGATCCGGCTCAAGAAAGAGGAGCGCTGTGCACACGCTTCACATTTGTGACAGCC 1560
Oy 1300 GTGAGCTTGCCTGATG-----GGTCTTCCCTTGGGAGATGATGCAATCTGCG 1353
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1361 ATCTGCTCTCGGAGCCGAGCAGCACTTCCCGCAGGACACAAATGTCCAGATTTGCGGG 1620
Oy 1354 TGGGG-----TGTTCAGAAACAGAAAGGTTCCCGCAGCTTCCTGATGCAAAAGTC 1407
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1621 TGGGGCACTTGTGATGAGAACGTGAGCGGCTACTCCAGCTTCCCTGCGGAGGCGCTGTC 1680

```

```

Oy 1408 AAGCTATGTCGCAACACTTTGTGCAACTCCCGCACTCTATGACCAATGATGATGAC 1467
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1681 CCCCTGTGTCGGAGCAAGATGTCAGACGCTCTGAGGTCTAGGCGCGGACATCAAGCCC 1740
Oy 1468 AGTATGATCTGTGCAAGAAATCTTCAGAAACCTGGGCAAGACACTGCGAGGTGACTCT 1527
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1741 AACATGCTCTGTGCGGCTACTT---CGACTCAAGTCCGACGCTTCCGAGGGGACTCA 1797
Oy 1528 GGAGGCCCCCTGACCTGTGAGAAAGGACGCGCACTTACGTATGAGGATGAGACTG 1587
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1798 GGGGGGCCCTGTGCTCGAGAAAGCGGTGCGCTTACCTTACGCAATCATCAATGCTGG 1857
Oy 1588 GCGCGAGTGTGAG-----AAGAGCCAGGCGGTCTACACCAAGTTACCAAAATCTCTG 1641
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1858 GGTGACGCGCTCGGGCGGTCCACAAGCCGGGGGTCTACACCGCGGCAACATATGTG 1917
Oy 1642 AATGTGATCAAGCCACCAT 1661
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1918 GACTGATCAAGACCGGAT 1937
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 8
US-09-879-792-33
; Sequence 33, Application US/09879792
; Patent No. US20020061850A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, Yonghong
; APPLICANT: Gedrich, Richard
; TITLE OF INVENTION: Regulation of Human Transmembrane Serine
; FILE OF INVENTION: Prolease
; FILE REFERENCE: 02973.00035
; CURRENT APPLICATION NUMBER: US/09/879,792
; PRIORITY FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/211,224
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US 60/283,353
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/283,648
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: PCT
; PRIOR FILING DATE: 2001-06-12
; NUMBER OF SEQ. ID NOS: 36
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ. ID NO 33
; LENGTH: 614
; TYPE: DNA
; ORGANISM: mouse
US-09-879-792-33

```

```

Query Match      4.3%; Score 72.8; DB 10; Length 614;
Best Local Similarity 57.2%; Pred. No. 2.9e-12;
Matches 175; Conservative 0; Mismatches 122; Indels 9; Gaps 2;

```

```

Oy 1381 TCCCGCAGCTCTGATGTCGCAAGTCAAGTGTGCAACACTTTGTGCAACTCCCGC 1440
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 195 TCTCCCTCTCCGAGAGGTTCAGTCAACTCATGTGATCAAGAAAGTGCATATGACTAC 254
Oy 1441 CAACCTTATGACCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 255 TTGGTCTATGACACTTACTTACCCCAAGAGATGATGTGCTCGGGGATCTACAGAGA--- 311
Oy 1501 GGGCAAGACACTGCGAGGCTGACTGTGAGGCGCCGCTGACCTGTGAGAAAGGAGCGCAC 1560
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 312 GGGAGAGACTCTTCCGAGGAGACAGTGSAGGACTCTCGTGTGAGCAGAAACAAATGCT 371
Oy 1561 TACTAGCTTATGGAGATGAGTGTGAGCTGGGCTCGAGTGTG-----AGAAGAGCGCAGG 1614
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 372 TGTGATCTGTGAGGTGTACACAGCTGGGGCAGAGCTGTGGCCAGAAACAAAGCTGTG 431
Oy 1615 GTCTACACCAAGTTTACCAAAATCTCTGATGATGATCAAAAGCCACCAATCAAAAGT 1674
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 432 GTGTACACCAAGTACAGAGTACTTCCCTGATTTACAGAAAGATGAGAGTGTAGGTA 491

```

Query	1675	GGCTTC	1680	
	1	111		
Db	492	CGATTTC	497	
RESULT 9				
US-09-969-271-6				
/ Sequence 6, Application US/09969271				
/ Patent No. US20020098179A1				
GENERAL INFORMATION:				
APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB)):				
APPLICANT: Pfizer Limited (GB and EP (GB) only)				
TITLE OF INVENTION: Pharmaceutical Combinations				
FILE REFERENCE: PCS10951APME				
CURRENT APPLICATION NUMBER: US/09/969, 271				
PRIORITY FILING DATE: 2001-10-01				
PRIOR APPLICATION NUMBER: GB 0025473.0				
NUMBER OF SEQ ID NOS: 7				
SOFTWARE: FastSeq for Windows Version 4.0				
SRO ID NO 6				
LENGTH: 1689				
TYPE: DNA				
ORGANISM: Homo sapiens				
US-09-969-271-6				
Query Match	4.3%	Score 72.2;	DB 10;	Length 1689;
Best Local Similarity	50.5%;	Pred. No. 8.4e-12;		
Matches 340;	Conservative 0;	Mismatches 288;	Indels 45;	Gaps 5
QY	1028	AGGGCAGCTTCTGTGTGTGGGGCGCTGATCCACCCCTGCTGGGTCTCACTGCTGCCACT	1087	
Db	1013	AGCGGTTCTGTGTGGGGCGATCACTATCAGCTCTCTGTGATTTCTGTGTGCCGCCACT	1072	
QY	1088	GCAACGA-----CATAAAAACGACATCTAAAGTGTGTCTAGGGGACCGAGACCTGA	1141	
Db	1073	GCTTCCAGGAGAGATTTCCGCCACCACTGACGTGTGATCTTGGGCAACAACTACCGG	1132	
QY	1142	AGAAAGAAATTTTCATGACGACAGTTTAGGTGTGCAAGAAATATTCAGTACACCCACT	1201	
Db	1133	TGTGTCCCTGTGGGAGAGAGACACAATTTGAAAGTGTGAAATATCATTTGTCCATAAGGAT	1192	
QY	1202	ACAAATGAAGAGATGAGATTTCCGCCATGATTTGCATTGTCTCAAGTTAAAGCCATGG	1261	
Db	1193	TGCAT-----GATGACACTTACGACAAAGATTTGGCTGTGACGCTGGAATGGGATT	1246	
QY	1262	ATGTGACACTGTGCTCTGAAATCCAAATACGTGAAGACTGTGTCTTGCCTGATGGTCTT	1321	
Db	1247	CGTCCCGGTGTGCCAAGAGAGAGAGGTGTGTCCGACTGTCTTCCCGGGGGGAC	1306	
QY	1322	TTTCCCT-----CTGGAGTGAATGCGACATCTCGCTGGGTGTTACAGAA	1369	
Db	1307	TGTAGCTGCCGAGCTGAGGAGATGTGAGCTTCCGGCTACGGCAAGCAATGAGGCTTTGT	1366	
QY	1370	CAGAAAGAGGTCGCCGACAGCTCTTGATGCCAAAGTCAAGCTGATTTGGCAACTTTGT	1429	
Db	1367	CTCTTCTTATTCGAGACGGCTGAAGAGAGGCTCATGTACAGCTGTACCCATCACCCCT	1426	
QY	1430	GCAACTCCCGCCACTCTATGACACCAATGATTTGATGACAGTATGATCTGTGAGGAATC	1489	
Db	1427	GCACTATCCAAATTTTACTTAAAGAAAGTACACGTAACCAATGCTGTGTGTGTGAGCA	1486	
QY	1490	TTTCAGAAACTGG-----GCAAGACCTCTCCAGGGTGTGACTGTGGAGCC	1534	
Db	1487	CTGGGAGCGGGGCCGCCAGCAAACTTCAAGACGCTCCAGAGCGATTTGGGAGGCC	1546	
QY	1535	CCCTGACCTGTGAAGAGAGGAGGACCTACTACGTCTATGGGATAGGAGCTGGGGCCGG	1594	
Db	1547	CCCTGTGTGTGTGAAGATGGCCGATGACTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1606	
QY	1595	AGTGTGAGAGAG-----CGAGGGGTCTACACCAATTACCAATTCCTGAATTGGA	1648	

	Dp	1607	GCTGGGACAGAGAATGTCCCGGGTGTACTACCAGAGTTTACAATTCTAAGTA	1666
Oy	1649	TCAAGCACCACAT	1661	
Dp	1667	TTTGTCACACAT	1679	
 RESULT 10 US-09-969-271-5 ; Sequence 5, Application US/09969271 ; Patent No. US20020098179A1 GENERAL INFORMATION: ; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB)); ; APPLICANT: Pfizer Limited (GB and EP (GB) only) ; TITLE OF INVENTION: Pharmaceutical Combinations ; FILE REFERENCE: PCS1095IAPME ; CURRENT APPLICATION NUMBER: US/09/969,271 ; PRIORITY FILING DATE: 2001-10-01 ; PRIOR APPLICATION NUMBER: GB 0025473.0 ; PRIORITY FILING DATE: 2000-10-17 ; NUMBER OF SEQ ID NOS: 7 ; SOFTWARE: FastSeq for Windows Version 4.0 ; SEQ ID NO: 5 ; LENGTH: 2519 ; TYPE: DNA ; ORGANISM: Homo sapiens US-09-969-271-5				
 Query Match 4.3%; Score 72.2; DB 10; Length 2519; Best Local Similarity 50.5%; Pred. No. 1,le-11; Matches 340; Conservative 0; Mismatches 288; Indels 45; Gaps 5				
Oy	1028	AGGGGCACCTTGTGGTGGGGGCGGTGATCACCCCTGCTGGGTGCTACCTGCCACT	1087	
Dp	1089	AGGGTTCCTGGGGGGGCGCATCTCATGCTCCTCGTGATTTCTCTGCCGCCACT	1148	
Oy	1088	GCACCGA-----CATAAACCAACATCTAAAGGTGTGTCGAAGGGACACGACCTGA	1141	
Dp	1149	GCCTCCAGAGAGAGTTCCTCCGCCCCACACACTGACGGGTATCTTGGGCGACAGACATACGGG	1208	
Oy	1142	AGAAAGAAATTTTCATGACGACAGCTTAGCGTGCGAAGATATTCAGTACAGCCACT	1201	
Dp	1209	TGGTCCCTGGGAGAGACGACCAATAATTTGAAAGTCGAAAATCATTTGTCCATAAGGAT	1268	
Oy	1202	ACAAATGAAGAATGATATTCGCCAACATGATATTCGATGTCACATTAAGCCACATGG	1261	
Dp	1269	TCGAT-----GATGACACTTATGACAAATGACATTGGCTGCGACGTGAATATGGCATTT	1322	
Oy	1262	ATGTGCTACTGTCTCTGAAATCCAAATACGTAAGACTGTGTCTTGCCGTAATGGCTCT	1321	
Dp	1323	CCTCCCGCTGGCCCAAGAGACGAGCGTGTGTCACACTGTCTCTTCCCCGGGGGAC	1382	
Oy	1322	TTTCCT-----CTGGGAGTGAATGCACATCTCTGCGCTGGGGTGTTCACAGAA	1369	
Dp	1383	TGAGCTGCCGCACTGACGAGAGTGTGACTCTCCGGCTACGGCAAGCATYGGAGCTTTGT	1442	
Oy	1370	CAGGAAGAAAGGTCCTCCGACACTTCTTGATGCCAAAGTCAACGTATTTGCCAACYTGT	1429	
Dp	1443	CTCTCTTCTTATTTGGAGCGSCTTAAGGAAGCTCATGTGACACTGTACCCATCCACGCT	1502	
Oy	1430	GCACTCTCCCGCACTCTATGACACATGATTTGATGACATGATGATCTGTGTGGGAATC	1489	
Dp	1503	GCACATCAACAATTTACTTAAACAGAACGTACCCGACAACAATGCTGTGTGTGGACAA	1562	
Oy	1490	TTTCAAGAACCTGG-----CCAAGACACTCCAGGGGTGAATCTGTGGAGGCC	1534	
Dp	1563	CTTGGAGGGGGGCCGCCCAAGCAAATCTTGACAGACGGCTGCCAGGGSGATTTGGAGGCC	1622	
Oy	1535	CCCTGACTGTGGAAGAGAGAGGACTTACTACGTCTATGGGATATGAGACTGGGGCTGG	1594	
Dp	1623	CCCTGATGTGTCTAAAGATGATGGCCGATGACTTTGTGGGGCATATCATCATCTGGGGCTGG	1682	
Oy	1595	AAGTGAAGAAAG-----CCAGGGGCTCTACACCACAAGTTACCAATTCCTGAAITGA	1648	

```
Db 1683 GCTGTGACAGAGATGTCCCGGTGTGTACACCAAGTTACCAACTTACCTGTGA 1742
Qy 1649 TCAAGGCCACCAT 1661
Db 1743 TTCTGTGACAAACAT 1755
```

RESULT 11

```
US-09-974-298-144
; Sequence 144, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Huei-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; PRIOR FILING DATE: 2001-10-04
; PRIORITY APPLICATION NUMBER: 60/238,331
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 144
; LENGTH: 2641
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020156263A1 1001470CBI
; NAME/KEY: unsure
; LOCATION: 2635
; OTHER INFORMATION: a, t, c, g, or other
US-09-974-298-144
```

```
Query Match 4.3%; Score 72.2; DB 9; Length 2641;
Best Local Similarity 50.5%; Pred. No. 1,le-11;
Matches 340; Conservative 0; Mismatches 288; Indels 45; Gaps 5;
```

```
Qy 1028 AGGGCCACTTCGTGTGGGGCGCTGATCCACCCCTGTGGGTGCTGCTGCTGCCCACT 1087
Db 1209 AGGGTTCTGTGGGGGCGATCTACTCATGCTCTGTGGATTTCTCTCCGCCCACT 1268
Qy 1088 GCACCGA-----CATMAAACACAGACATCTAAGGTGTGTAGGGAGCCAGCACTGA 1141
Db 1269 GCTTCCAGAGAGGTTTCCGCCCACTGACGGTATCTTGGGAGCAACATACCGGG 1328
Qy 1142 AGAAAGAAATTTATGACGAGCTTTAGGGTGCAGAGATATTTCAAGTACAGCCACT 1201
Db 1329 TGGTCCCTGGCGAGGAGGAGCAAAATTTGAAGTCGMAAAATACATTTCCATGAAGAA 1388
Qy 1202 ACAATGAAGAGATGAGATTTCCCAATATGATTTGCTTGAAGTTAAGCCAGTGG 1261
Db 1389 TCGAT-----GATGACACTTACGACATGACATTTGCGTGTGTGAGCTGAATTCGAT 1442
Qy 1262 ATGTGACTGTGCTGTAGATTCAAATACCTGAGAGCTGTGTGCTTGGCTGATGGTCT 1321
Db 1443 CGTCCGCTGTCCCGAGGAGAGCGCTGTGCGCATGTGTGCTTCCCGCGGAGCC 1502
Qy 1322 TTCCCT-----CTGGAGTGTAGTGCACATCTCTGTGGTGGGGGTGTTACAGAAA 1369
Db 1503 TTAGAGTGTCCGAGATGACGCGAGTGTGAGTCTCCGCGCTACCGGAGAGATGAGCC 1562
Qy 1370 CAGGAAAGGCTCCCGCAGCTCTGTGATGCCAAAGTCAAGCTGATTTGCCAACA 1429
Db 1563 CTCTCTTTTATTCGAGGAGGCTGAAGAGCTCATGTGATGATACCATTCAGCGCT 1622
Qy 1430 GCACTTCCCGCACTCTATGATACCATGATGTGATGATGATGATGATGATGATGAT 1489
Db 1623 GCACATCACAACATTTACTTAAGAAAGTCAACGACATGATGTGTGTGTGTGTGTGT 1682
Qy 1490 TTGAGAAACTGG-----GCAAGACACTGCGAGGAGTACTCTGTGAGGCC 1534
Db 1683 CTGCGAGCGCGGCGCCCGAGCAAACTTGTGACGACGCTGCCAGGGGATTTGAGGAGCC 1742
```

```
Qy 1535 CCCTGACCTGTGAGAGAGCGCACTACTAGCTATGATGATAGTACCTGGGCTTCGG 1594
Db 1743 CCCTGGTGTGTCTGACACATGCGCGATGACTTTGGTGATCATGACCTGGGCTTCGG 1802
Qy 1595 AGTGTGAGAGAG-----CGAGGGTCTACACCAAGTTACCAAAATTCCTGAATTTGA 1648
Db 1803 GCTGTGACAGAGAGATGTCCCGGTGTGTGTACACCAAGTTACCAACTACTAGACTGA 1862
Qy 1649 TCAAGGCCACCAT 1661
Db 1863 TTCTGTGACAAACAT 1875
```

RESULT 12

```
US-09-888-615-45
; Sequence 45, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: DLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEL, SEAN
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; PRIOR FILING DATE: 2001-06-26
; PRIORITY APPLICATION NUMBER: 60/214,047
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 1614
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-888-615-45
```

```
Query Match 3.9%; Score 66.4; DB 10; Length 1614;
Best Local Similarity 55.9%; Pred. No. 5,2e-10;
Matches 171; Conservative 0; Mismatches 126; Indels 9; Gaps 2;
```

```
Qy 1381 TCCCGCCAGCTCTCTGTGAGCAAAAGTCAAGCTGATTTGCCAACCTTTGCCAATCCGC 1440
Db 1300 TCCCGCTTCTCCGAGGAGTGCAGGTCAATCTCATGCTCAAGAAATGCAATGACTAC 1359
Qy 1441 CAACCTATGACCAATGATGTATGATGATGATGATGATGATGATGATGATGATGAT 1500
Db 1360 TTGTGTATGACATTTACTTACCCCAAGATGATGTGTGTGTGTGTGTGTGTGTGTGT 1416
Qy 1501 GGGCAAGACCTTCCAGGCTGTCTGTGAGGCCCTTGACCTGTGAGAGAGCGCAC 1560
Db 1417 GGCAGAGACTCTTCCAGGAGACAGCGGGGCGCTCTGTGTGTGTGTGTGTGTGTGTGT 1476
Qy 1561 TACTATGCTATGAGTGTGAGCTGTGGGCTGGAGTGTGAGAG-----GGCAGG 1614
Db 1477 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1536
Qy 1615 GTCTACACCAAGTTACAAATTTCTGAATTTGATCAAGCCACCATCAAAAGTGAAGT 1674
Db 1537 GTGTACACCAAGTGAAGAGTGTCTTCCCTGTGATTTACAGCAAGATGAGAGCGAG 1596
Qy 1675 GGCCTC 1680
Db 1597 CGATTTC 1602
```

RESULT 13

```
US-09-879-792-11
; Sequence 11, Application US/09879792
; Patent No. US20020061850A1
; GENERAL INFORMATION:
```


GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Geritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James.
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OR INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P26301C27
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: US/09/978,697
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27

PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545

1	PRIOR FILING DATE: 1998-04-29
2	PRIOR APPLICATION NUMBER: 60/083554
3	PRIOR FILING DATE: 1998-04-29
4	PRIOR APPLICATION NUMBER: 60/083558
5	PRIOR FILING DATE: 1998-04-29
6	PRIOR APPLICATION NUMBER: 60/083559
7	PRIOR FILING DATE: 1998-04-29
8	PRIOR APPLICATION NUMBER: 60/083500
9	PRIOR FILING DATE: 1998-04-29
10	PRIOR APPLICATION NUMBER: 60/083742
11	PRIOR FILING DATE: 1998-04-30
12	PRIOR APPLICATION NUMBER: 60/084366
13	PRIOR FILING DATE: 1998-05-05
14	PRIOR APPLICATION NUMBER: 60/084414
15	PRIOR FILING DATE: 1998-05-06
16	PRIOR APPLICATION NUMBER: 60/084441
17	PRIOR FILING DATE: 1998-05-06
18	PRIOR APPLICATION NUMBER: 60/084637
19	PRIOR FILING DATE: 1998-05-07
20	PRIOR APPLICATION NUMBER: 60/084639
21	PRIOR FILING DATE: 1998-05-07
22	PRIOR APPLICATION NUMBER: 60/084640
23	PRIOR FILING DATE: 1998-05-07
24	PRIOR APPLICATION NUMBER: 60/084598
25	PRIOR FILING DATE: 1998-05-07
26	PRIOR APPLICATION NUMBER: 60/084600
27	PRIOR FILING DATE: 1998-5--07
28	PRIOR APPLICATION NUMBER: 60/084627
29	PRIOR FILING DATE: 1998-05-07
30	PRIOR APPLICATION NUMBER: 60/084643
31	PRIOR FILING DATE: 1998-05-07
32	PRIOR APPLICATION NUMBER: 60/085339
33	PRIOR FILING DATE: 1998-05-13
34	PRIOR APPLICATION NUMBER: 60/085338
35	PRIOR FILING DATE: 1998-05-13
36	PRIOR APPLICATION NUMBER: 60/085323
37	PRIOR FILING DATE: 1998-05-13
38	PRIOR APPLICATION NUMBER: 60/085582
39	PRIOR FILING DATE: 1998-05-15
40	PRIOR APPLICATION NUMBER: 60/085700
41	PRIOR FILING DATE: 1998-05-15
42	PRIOR APPLICATION NUMBER: 60/085689
43	PRIOR FILING DATE: 1998-05-15
44	PRIOR APPLICATION NUMBER: 60/085579
45	PRIOR FILING DATE: 1998-05-15
46	PRIOR APPLICATION NUMBER: 60/085580
47	PRIOR FILING DATE: 1998-05-15
48	PRIOR APPLICATION NUMBER: 60/085573
49	PRIOR FILING DATE: 1998-05-15
50	PRIOR APPLICATION NUMBER: 60/085704
51	PRIOR FILING DATE: 1998-05-15
52	PRIOR APPLICATION NUMBER: 60/085697

[illegible]

Oy	1547	AGAGAGCAGCGC	ACTACTACGCTGTATGGGATAGTCACTGGGCGCTTGAGATG-----TG	1600
Db	1345	AAGAGAGGAGGCTGTGGCAAGTTAGTGGGAGCAGCAGCTTGGCATGGGCTGGCGAGAG		1404
Oy	1601	AGAGAGGCGCAGGGGTGTACACCCAGTAGTACCAAAATTCCTAATTTGGATCAAGCCACCA		1660
Db	1405	TGAAACAAGCTGTGGGGTGTACACCCGCTGTCACTCCTTCTCTGAGATGGATGCAGAGACAGA		1464
Oy	1661	TCCAAAGTGA	1670	
Db	1465	TGGAGAGAGA	1474	

```
Search completed: March 6, 2003, 23:07:10
Job time : 137.5 secs
```


GenCore version 5.1.4_p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 10:42:18 : Search time 342.5 Seconds
(without alignments)
11066.024 Million cell updates/sec

Title: US-09-912-559-2

Perfect score: 1683

Sequence: 1 atgttgcgcagatgtctga.....aaagtgaagtgccttctaa 1683

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

```

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1683	100.0	1683	24	AAL45697
2	1679.8	99.8	1683	24	AAL45696
3	1679.8	99.8	3008	24	ABN95170
4	1675	99.5	3623	21	AACT6693
5	526.4	31.3	617	21	AAA44763
6	241	14.3	397	24	AB199281
7	120.2	7.1	451	22	ABA58323
8	120.2	7.1	451	22	AAK06420
9	120.2	7.1	451	22	AAK32096

C	10	120.2	7.1	451	22	AA137947	Probe #6633 used t
C	11	120.2	7.1	451	24	AB506863	Human genome-deriv
C	12	117	7.0	117	22	ABA70919	Human foetal liver
C	13	117	7.0	117	22	AAK19188	Human brain expres
C	14	117	7.0	117	22	AAK45153	Human bone marrow
C	15	117	7.0	117	22	AAK45106	Probe #19792 used
C	16	117	7.0	117	22	AB519403	Human genome-deriv
C	17	114.4	6.8	2033	15	AA063951	Hepatocyte growth
C	18	114.4	6.8	2036	24	ABN95114	Gene #1612 used to
C	19	114.4	6.8	2036	24	ABL65242	lung cancer relate
C	20	93	5.5	970	15	AA063945	Gene compising HG
C	21	87.6	5.2	2252	11	AA005673	v-PA.alpha1. Desm
C	22	87.4	5.2	2512	12	AA012867	JM1-229 cell line
C	23	86.8	5.2	1323	8	AA012867	Modified tissue pl
C	24	85.2	5.0	1422	8	AA070991	v-PA.alpha2. Desm
C	25	84.8	5.0	1422	11	AA005675	Modified tissue pl
C	26	79.6	4.7	2257	11	AA005674	v-PA.alpha2. Desm
C	27	78	4.6	1620	11	AA000543	Vampire bat plasml
C	28	77.2	4.6	919	18	AA090048	Chicken urokinase
C	29	75.4	4.5	1689	10	AA091217	Mutated CDNA codin
C	30	73.8	4.4	2097	12	AA012071	T-PA with -ve char
C	31	73.8	4.4	2100	12	AA012074	T-PA variant havin
C	32	72.8	4.3	614	24	AA027741	Human full length
C	33	72.2	4.3	329	15	AA079004	Gene compising HG
C	34	72.2	4.3	1047	14	AA040658	tpa-2 CDNA. Synth
C	35	72.2	4.3	1065	18	AA040658	mt-PA6-E. Homo sa
C	36	72.2	4.3	1065	18	AA040658	DNA encoding t-PA
C	37	72.2	4.3	1068	14	AA053322	Human truncated tP
C	38	72.2	4.3	1068	14	AA053322	Human truncated tP
C	39	72.2	4.3	1068	14	AA040657	tpa-1 CDNA. Synth
C	40	72.2	4.3	1068	14	AA040657	tpa-6 CDNA. Synth
C	41	72.2	4.3	1068	14	AA040660	tpa-11 CDNA. Synth
C	42	72.2	4.3	1068	14	AA040669	tpa-12 CDNA. Synth
C	43	72.2	4.3	1068	14	AA040671	tpa-13 CDNA. Synth
C	44	72.2	4.3	1068	14	AA040673	tpa-14 CDNA. Synth
C	45	72.2	4.3	1068	14	AA040677	tpa-16 CDNA. Synth

ALIGNMENTS

RESULT 1	
AAL45697	standard: DNA: 1683 BP.
ID	AA145697 standard: DNA: 1683 BP.
AC	AA145697:
XX	
DT	13-JUN-2002 (first entry)
DE	Human blood coagulation factor VII activating protease mutant DNA.
XX	
KW	Human: blood coagulation factor VII activating protease: FSAP.
KW	Single-chain plasminogen activator: bleeding disorder: haematological:
KW	haemostatic; mutant; gene; ds.
XX	
OS	Homo sapiens.
OS	Synthetic.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..1683
FT	/*tag= a
FT	/product= "mutant FSAP"
PN	EP1182258-A1.
XX	
PD	27-FEB-2002.
XX	
PF	05-JUL-2001: 2001EP-0115691.
XX	
PR	26-JUL-2000: 2000DE-1036641.
PR	10-OCT-2000: 2000DE-1050040.
PR	21-OCT-2000: 2000DE-1052319.
PR	12-APR-2001: 2001DE-1018706.

XX (AVET) AVENTIS BEHRING GMBH.
 XX
 PI Roemisch J, Stoeck H, Feussner A, Lang W, Weimer T, Becker M;
 PI Nerlich C, Muth-Naumann G;
 XX
 DR MPI: 2002-270939/32.
 DR P-PSDB: AMO17145.

XX New nucleic acid encoding mutant factor 7 activating protease, useful
 PT for diagnosis, treatment and prevention of coagulation disorders, also
 PT related protein and antibodies -
 XX
 PS Disclosure: Page 16-17; 27p: German.

XX The present invention relates to a mutant of the DNA sequence encoding
 CC the protease (F5A) that activates blood coagulation factor VII (FVII)
 CC and single-chain plasminogen activator, where at least one of the base
 CC changes G to C at nucleotide 1177 and G to A at nucleotide 1601 is
 CC present. The mutant sequences can be used in the treatment and prevention
 CC of bleeding disorders associated with inherited or acquired defects of
 CC blood coagulation factors V, VIII, IX, X, XI, XII, von Willebrand's
 CC factor and/or with antibodies against any of these proteins. The present
 CC sequence is the mutant human F5A coding sequence.

SQ Sequence 1683 BP; 441 A; 437 C; 435 G; 370 T; 0 other;

Query Match 100.0%; Score 1683; DB 24; Length 1683;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1683; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGTTGGCAGAGATGTGATTCATGCTGCTGTTGATGATGCTGGTGGGAAAGACA 60
 DB 1 ATGTTGGCAGAGATGTGATTCATGCTGCTGTTGATGATGCTGGTGGGAAAGACA 60
 OY 61 GCCTGTGGGTTCTCCCTGATGCTTTATTGGAAGCCTGAGCCAGACTGGACCCCTGAC 120
 DB 61 GCCTGTGGGTTCTCCCTGATGCTTTATTGGAAGCCTGAGCCAGACTGGACCCCTGAC 120
 OY 121 CAGTTGATTTACAGTACGAGATTTAATCAGAGAAGAACACACATACACACTTACC 180
 DB 121 CAGTTGATTTACAGTACGAGATTTAATCAGAGAAGAACACACATACACACTTACC 180
 OY 181 CATGCTGAAGATCTGACTGATCTACTAGTACGAGCAAGCTGATCCAGCCCAAC 240
 DB 181 CATGCTGAAGATCTGACTGATCTACTAGTACGAGCAAGCTGATCCAGCCCAAC 240
 OY 241 CCTGTGAACAGGTGGGAGTCTGCTGCTGATGAGGACACCTTACATGACACTGCTG 300
 DB 241 CCTGTGAACAGGTGGGAGTCTGCTGCTGATGAGGACACCTTACATGACACTGCTG 300
 OY 301 GCTCTTTCTCTGGGAATTAAGTTCAGAAAGTGAATACGTGCAAGCAACCATGT 360
 DB 301 GCTCTTTCTCTGGGAATTAAGTTCAGAAAGTGAATACGTGCAAGCAACCATGT 360
 OY 361 GGCCTGGGCAATGTCTATTAACCCAGAGTCTCCCTACTACCCCTGTGCTTAACAC 420
 DB 361 GGCCTGGGCAATGTCTATTAACCCAGAGTCTCCCTACTACCCCTGTGCTTAACAC 420
 OY 421 CCTTACACAGTCCAGCTGCTCCCAAGTGTCTCTGATGACGACCAACCCCTGACG 480
 DB 421 CCTTACACAGTCCAGCTGCTCCCAAGTGTCTCTGATGACGACCAACCCCTGACG 480
 OY 481 AATGGGGCTACCTGCTCCGGCATTAAGCGAGATCAAGTTCACTGTGCTTCCGAC 540
 DB 481 AATGGGGCTACCTGCTCCGGCATTAAGCGAGATCAAGTTCACTGTGCTTCCGAC 540
 OY 541 CAGTTCAAGGGAATTTCTGTGAATAGTCTGATGACTCTATGTGGCGATGCTAC 600
 DB 541 CAGTTCAAGGGAATTTCTGTGAATAGTCTGATGACTCTATGTGGCGATGCTAC 600
 OY 601 TCTTACCGGAGAAATGATAGACAGTCAACGACGATGCTGCTTACTGGAAGTCC 660
 DB 601 TCTTACCGGAGAAATGATAGACAGTCAACGACGATGCTGCTTACTGGAAGTCC 660

DB 601 TCTTACCGGAGAAATGATAGACAGTCAACGACGATGCTGCTTACTGGAAGTCC 660
 OY 661 CACCTCTCTTTCAGAGAAATTTACACATGTTTATGAGAGATGCTGAACCCATGGATT 720
 DB 661 CACCTCTCTTTCAGAGAAATTTACACATGTTTATGAGAGATGCTGAACCCATGGATT 720
 OY 721 GGGGAACACAAATTTCTGAGAAACCCAGATGCGAGCAAAAGCCCTGCTGCTTATTAA 780
 DB 721 GGGGAACACAAATTTCTGAGAAACCCAGATGCGAGCAAAAGCCCTGCTGCTTATTAA 780
 OY 781 GTTACCAATGACAAAGTGAATTTGGAATTTACTGATGCTCAGCTGCTCAGCCAGAC 840
 DB 781 GTTACCAATGACAAAGTGAATTTGGAATTTACTGATGCTCAGCTGCTCAGCCAGAC 840
 OY 841 GTTGCTTACCCAGAGAAAGCCCATGAGCCATCAACCAAGCTTCCGGGTTTGAAGTCC 900
 DB 841 GTTGCTTACCCAGAGAAAGCCCATGAGCCATCAACCAAGCTTCCGGGTTTGAAGTCC 900
 OY 901 TGTGAAAGACTGAGATAGCAGAGAGAGATCAAGAGATCTATGAGGCTTTAAGAGC 960
 DB 901 TGTGAAAGACTGAGATAGCAGAGAGAGATCAAGAGATCTATGAGGCTTTAAGAGC 960
 OY 961 ACGGCGGAGACACCCCATGAGGAGGCTCCCTCAGTCTGCTGCTGCTGACCATCTCC 1020
 DB 961 ACGGCGGAGACACCCCATGAGGAGGCTCCCTCAGTCTGCTGCTGCTGACCATCTCC 1020
 OY 1021 ATGCCCCAGGCGCACTTCTGTGTGGGCGCTGATCCACCCCTGCTGAGTCTGCTGCT 1080
 DB 1021 ATGCCCCAGGCGCACTTCTGTGTGGGCGCTGATCCACCCCTGCTGAGTCTGCTGCT 1080
 OY 1081 GCCCAGTGCACGACATTAAGAAACAGACATCTAAAGTGTGCTGAGGAGACAGAGCTG 1140
 DB 1081 GCCCAGTGCACGACATTAAGAAACAGACATCTAAAGTGTGCTGAGGAGACAGAGCTG 1140
 OY 1141 AAGAAAGAAATTTCTATGAGACAGAGCTTTAGGCTCAGAGATTTTCAAGTACAGCAC 1200
 DB 1141 AAGAAAGAAATTTCTATGAGACAGAGCTTTAGGCTCAGAGATTTTCAAGTACAGCAC 1200
 OY 1201 TACAAATGAAGAGATGAGATTTCCCAATGATTTGCAATTCCTCAAGTTTAAAGCCAGT 1260
 DB 1201 TACAAATGAAGAGATGAGATTTCCCAATGATTTGCAATTCCTCAAGTTTAAAGCCAGT 1260
 OY 1261 GATGTCACCTGCTGCTGATTTCAATTTAGTGAAGACCTGCTGCTGCTGAGGCTCC 1320
 DB 1261 GATGTCACCTGCTGCTGATTTCAATTTAGTGAAGACCTGCTGCTGCTGAGGCTCC 1320
 OY 1321 TTTCCCTGAGGAGTGAAGTGCACATCTGAGCTGGGAGTTTACAGAAACAGAAAGGG 1380
 DB 1321 TTTCCCTGAGGAGTGAAGTGCACATCTGAGCTGGGAGTTTACAGAAACAGAAAGGG 1380
 OY 1381 TCCCGCAGCTCTGGATGCCAAGCTCAAGCTGATGCGCAACACTTGTGCAACTCCGCG 1440
 DB 1381 TCCCGCAGCTCTGGATGCCAAGCTCAAGCTGATGCGCAACACTTGTGCAACTCCGCG 1440
 OY 1441 CAACCTATGACACATGATGATGAGACAGATGATCTGTGACAGAAATCTTCAGAAACCT 1500
 DB 1441 CAACCTATGACACATGATGATGAGACAGATGATCTGTGACAGAAATCTTCAGAAACCT 1500
 OY 1501 GGGGACACACCTGCGAGGAGTACTGTGAGAGCCCTGACCTGTGAAGAGAGAGGACCC 1560
 DB 1501 GGGGACACACCTGCGAGGAGTACTGTGAGAGCCCTGACCTGTGAAGAGAGAGGACCC 1560
 OY 1561 TACTAGCTATGAGATGAGTGAAGTGGGGCTGAGTGTGAAGAGAGGCGAGGCTTAC 1620
 DB 1561 TACTAGCTATGAGATGAGTGAAGTGGGGCTGAGTGTGAAGAGAGGCGAGGCTTAC 1620
 OY 1621 ACCCAATTAACCAATTTCTGATTTGATCAAGCCATCAAAAGTGAAGAGGCTTC 1680
 DB 1621 ACCCAATTAACCAATTTCTGATTTGATCAAGCCATCAAAAGTGAAGAGGCTTC 1680
 OY 1681 TAA 1683
 DB 1681 TAA 1683

RESULT 2
ID AAL45696 standard; DNA; 1683 BP.
XX AAL45696;
AC AAL45696;
XX 13-JUN-2002 (first entry)
DE Human blood coagulation factor VII activating protease DNA.
XX
XX Human; blood coagulation factor VII activating protease; FSAP;
KW Single-chain plasminogen activator; bleeding disorder; haematological;
KW haemostatic; gene; ds.
XX
XX Homo sapiens.
FH Key Location/Qualifiers
FT CDS 1..1683
FT /tag= a
FT /product= "FSAP"
XX
XX EP1182258-A1.
XX
XX 27-FEB-2002.
XX
XX 05-JUL-2001; 2001EP-0115691.
XX
XX 26-JUL-2000; 2000DE-1036641.
XX 10-OCT-2000; 2000DE-1050040.
XX 21-OCT-2000; 2000DE-1052319.
XX 12-APR-2001; 2001DE-1018706.
XX
XX (AVET) AVENTIS BEHRING GMBH.
XX
XX Roemisch J, Steehr H, Feussner A, Lang W, Weimer T, Becker M;
XX Nerlich C, Muth-Naumann G;
XX P-PSDB: AAO17144.
XX
XX WPI: 2002-270939/32.
XX
XX New nucleic acid encoding mutant factor 7 activating protease, useful
PT for diagnosis, treatment and prevention of coagulation disorders, also
PT related protein and antibodies
XX
XX Claim 2; Page 15-16; 27pp; German.
XX
XX The present invention relates to a mutant of the DNA sequence encoding
CC the protease (FSAP) that activates blood coagulation factor VII (FVII)
CC and single-chain plasminogen activator, where at least one of the base
CC changes G to C at nucleotide 1177 and G to A at nucleotide 1601 is
CC present. The mutant sequences can be used in the treatment and prevention
CC of bleeding disorders associated with inherited or acquired defects of
CC blood coagulation factors V, VIII, IX, X, XI, XII, von Willebrand's
CC factor and/or with antibodies against any of these proteins. The present
CC sequence is the human FSAP coding sequence.
XX
XX Sequence 1683 BP; 440 A; 436 C; 437 G; 370 T; 0 other;
SO
Query Match 99.8%; Score 1679.8; DB 24; Length 1683;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1681; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 121 CAGTATGATTACAGCTACGAGATTATATACAGGAAGAGAACACCAGTACACTTACC 180
Qy 181 CANGTGAGAAATCTGACTGGTACTACACTGAGACCAGCTGATCATGACACCCAAC 240
Db 181 CATGCTGAGAAATCTGACTGGTACTACACTGAGACCAGCTGATCATGACACCCAAC 240
Qy 241 CCCTGTGAACACGCTGGGACACTGCTCCATGAGAGCACCTTCACATGACAGCTGCTG 300
Db 241 CCCTGTGAACACGCTGGGACACTGCTCCATGAGAGCACCTTCACATGACAGCTGCTG 300
Qy 301 GCTCCTTTCTGCGGAATTAAGTTCAGAAAAGTGCAAAATACGTGCAAGACCAACCTATG 360
Db 301 GCTCCTTTCTGCGGAATTAAGTTCAGAAAAGTGCAAAATACGTGCAAGACCAACCTATG 360
Qy 361 GGCCGGGGCCCAATGTCTCATTAACCCAGATCTCCCTACTACCGCTGTGTCTGTAAACAC 420
Db 361 GGCCGGGGCCCAATGTCTCATTAACCCAGATCTCCCTACTACCGCTGTGTCTGTAAACAC 420
Qy 421 CCTTACACAGTCCCGACGTCTCCCAAGTGGTTCGTATAGAGGCAACCCCTGCGCAG 480
Db 421 CCTTACACAGTCCCGACGTCTCCCAAGTGGTTCGTATAGAGGCAACCCCTGCGCAG 480
Qy 481 AATGGGGCTACCTGCTCCCGCATTAAGCGAGATCCCAAGTTGACCTGTGCTGCCGAC 540
Db 481 AATGGGGCTACCTGCTCCCGCATTAAGCGAGATCCCAAGTTGACCTGTGCTGCCGAC 540
Qy 541 CAGTTCAAGGGGAAATTCGTGGAATTAAGTTCTGATGACCTGATGTTGGGATGAGCTAC 600
Db 541 CAGTTCAAGGGGAAATTCGTGGAATTAAGTTCTGATGACCTGATGTTGGGATGAGCTAC 600
Qy 601 TCTTACCGAGGGAATTAAGTATAGCACTCAACAGCATGCGCTTACTGGAAGCTAC 660
Db 601 TCTTACCGAGGGAATTAAGTATAGCACTCAACAGCATGCGCTTACTGGAAGCTAC 660
Qy 661 CACCTCCTCTGAGAGAGATTAACACATGTTATGAGATGCTGGAACCCATGGGATT 720
Db 661 CACCTCCTCTGAGAGAGATTAACACATGTTATGAGATGCTGGAACCCATGGGATT 720
Qy 721 GGGGACACCAATTTCTGCAAAACCAGATCGGACGAAAAAGCCCTGCTGTTATTAAA 780
Db 721 GGGGACACCAATTTCTGCAAAACCAGATCGGACGAAAAAGCCCTGCTGTTATTAAA 780
Qy 781 GTTACCAATGACAAGTGAATAGGAATACGTGATGCTCAGCCCTGCTCAGCCAGAC 840
Db 781 GTTACCAATGACAAGTGAATAGGAATACGTGATGCTCAGCCCTGCTCAGCCAGAC 840
Qy 841 GTTGCCTACCCAGAGGAAAGCCCACTGAGCCATCAACCAAGCTTCGGGGTTTGACTCC 900
Db 841 GTTGCCTACCCAGAGGAAAGCCCACTGAGCCATCAACCAAGCTTCGGGGTTTGACTCC 900
Qy 901 TGTGGAAGACTAGATTAACAGAGGAAATCAAGAGAAATCATGAGAGCTTTAAAGAC 960
Db 901 TGTGGAAGACTAGATTAACAGAGGAAATCAAGAGAAATCATGAGAGCTTTAAAGAC 960
Qy 961 ACGGCGGGCAAGACACCATGAGGAGGCTCCCTCCAGTCTCGCTGCTTGACATCTCC 1020
Db 961 ACGGCGGGCAAGACACCATGAGGAGGCTCCCTCCAGTCTCGCTGCTTGACATCTCC 1020
Qy 1021 ATGCCCAAGGGCCATCTTCTGTGTGGGGCTGATCCACCCCTGCTGGGTGCTACGTCT 1080
Db 1021 ATGCCCAAGGGCCATCTTCTGTGTGGGGCTGATCCACCCCTGCTGGGTGCTACGTCT 1080
Qy 1081 GCCCATGCAACCCAGATTAATAAACCCAGACATTAAGAGTGTGTAGGGAGCCAGACTGT 1140
Db 1081 GCCCATGCAACCCAGATTAATAAACCCAGACATTAAGAGTGTGTAGGGAGCCAGACTGT 1140
Qy 1141 AAGGAAGAAGAAATTTATAGAGAGAGCTTTAGGGTCAAGAAGATATTCAAGTACAGCCAC 1200
Db 1141 AAGGAAGAAGAAATTTATAGAGAGAGCTTTAGGGTCAAGAAGATATTCAAGTACAGCCAC 1200
Qy 1201 TACAAATGAAGAGATGAGATTCCTCCCAATGATATTGCAATTCCTCAAGTTAAAGCAGTGT 1260
Db 1201 TACAAATGAAGAGATGAGATTCCTCCCAATGATATTGCAATTCCTCAAGTTAAAGCAGTGT 1260

Db 97 ATGTTGCCAGGATGCTGATCTCCATGTTCTGCTTTATGGCTCGGTGGGAAAAGACA 156
Qy 61 GCCGTGGGTTCTCCGATGTCCTTTATTTGAAAACCCCTGAGCCAGACTGAGACCCCTGAC 120
Db 157 GCGTGTGGGTTCTCCGATGTCCTTTATTTGAAAACCCCTGAGCCAGACTGAGACCCCTGAC 216
Qy 121 CAGTGTGATTTACACTGAGGATTTATATCAGGAAGGAACCCATGACACTTACC 180
Db 217 CAGTGTGATTTACACTGAGGATTTATATCAGGAAGGAACCCATGACACTTACC 276
Qy 181 CATGCTGGAATTCCTGACTGCTACTACACTGAGACCAAGCTGATCCAGCCCAAC 240
Db 277 CACGCTGAGAACTCTGACTGCTACTACACTGAGACCAAGCTGATCCAGCCCAAC 336
Qy 241 CCGTGTGAACGGGTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db 337 CCGTGTGAACGGGTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 396
Qy 301 GCTCCTTTCTGTGGGAATAGTGTGCAAAAATGCAAAATACGTGCAAGCAACCCATGT 360
Db 397 GCTCCTTTCTGTGGGAATAGTGTGCAAAAATGCAAAATACGTGCAAGCAACCCATGT 456
Qy 361 GCGCGGGGCAATGTCATTTACCCAGAGTCCCTCTACCTGCTGCTGCTGCTGCTGCTG 420
Db 457 GCGCGGGGCAATGTCATTTACCCAGAGTCCCTCTACCTGCTGCTGCTGCTGCTGCTG 516
Qy 421 CCTTACACAGGTCCAGCTGCTCCCAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 517 CCTTACACAGGTCCAGCTGCTCCCAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 576
Qy 481 AATGGGGGTACTGCTCCGGCATAGCGAGATCCCAATTCACCTGCTGCTGCTGCTGCTG 540
Db 577 AATGGGGGTACTGCTCCGGCATAGCGAGATCCCAATTCACCTGCTGCTGCTGCTGCTG 636
Qy 541 CAGTTCAGGGGAAAATTCGTGTAATAGTTCGTGATGCTGCTGCTGCTGCTGCTGCTG 600
Db 637 CAGTTCAGGGGAAAATTCGTGTAATAGTTCGTGATGCTGCTGCTGCTGCTGCTGCTG 696
Qy 601 TCTTACCGAGGAAAATGATAGGATAGACAGTCAACAGCTGCTGCTGCTGCTGCTGCTG 660
Db 697 TCTTACCGAGGAAAATGATAGGATAGACAGTCAACAGCTGCTGCTGCTGCTGCTGCTG 756
Qy 661 CACCTCTCTTTCGAGAGAAATTTACAACTGTTTATGAGAGTCTGTAACCCATGGGAT 720
Db 757 CACCTCTCTTTCGAGAGAAATTTACAACTGTTTATGAGAGTCTGTAACCCATGGGAT 816
Qy 721 GGGGAACACAAATTTCTGACAGAAACCCAGATGCGAGAAAAGCCCTGCTGCTTATTA 780
Db 817 GGGGAACACAAATTTCTGACAGAAACCCAGATGCGAGAAAAGCCCTGCTGCTTATTA 876
Qy 781 GTTACCATGACAAAGGGAATGGGAATGCTGATGTCCTAGCCTGCTAGCCCAAGAC 840
Db 877 GTTACCATGACAAAGGGAATGGGAATGCTGATGTCCTAGCCTGCTAGCCCAAGAC 936
Qy 841 GTTACCATGACAAAGGGAATGGGAATGCTGATGTCCTAGCCTGCTAGCCCAAGAC 900
Db 937 GTTACCATGACAAAGGGAATGGGAATGCTGATGTCCTAGCCTGCTAGCCCAAGAC 996
Qy 901 TGTGGAAGACCTGAGATGACAGAGAGAAAGATCAAGAAATCTATGAGGCTTTAAGAC 960
Db 997 TGTGGAAGACCTGAGATGACAGAGAGAAAGATCAAGAAATCTATGAGGCTTTAAGAC 1056
Qy 961 ACAGGCGGAGACACCCATGAGAGGCTCCCTGACAGTCTGCTGCTGCTGCTGCTGCTGCT 1020
Db 1057 ACAGGCGGAGACACCCATGAGAGGCTCCCTGACAGTCTGCTGCTGCTGCTGCTGCTGCT 1116
Qy 1021 ATGCCCCAGGGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Db 1117 ATGCCCCAGGGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1176
Qy 1081 GCGCACTGACGAGATTAAGGAGTAAAGGAGTAAAGGAGTAAAGGAGTAAAGGAGTAA 1140
Db 1177 GCGCACTGACGAGATTAAGGAGTAAAGGAGTAAAGGAGTAAAGGAGTAAAGGAGTAA 1236

Qy 1141 AAGAAGAAGAAATTTTCATGAGCAGAGCTTTAGGGTGCAGAGATATTTCAAGTACAGCCAC 1200
Db 1237 AAGAAGAAGAAATTTTCATGAGCAGAGCTTTAGGGTGCAGAGATATTTCAAGTACAGCCAC 1296
Qy 1201 TACATGAAGAAGATGAGATTTCCCAATGATATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Db 1297 TACATGAAGAAGATGAGATTTCCCAATGATATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1356
Qy 1261 GATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
Db 1357 GATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1416
Qy 1321 TTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
Db 1417 TTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1476
Qy 1381 TCCCGCAGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
Db 1477 TCCCGCAGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1536
Qy 1441 CAACCTGATGACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
Db 1537 CAACCTGATGACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1596
Qy 1501 GGGCAGACACCTGCGAGGCTGCTGCGAGGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
Db 1597 GGGCAGACACCTGCGAGGCTGCTGCGAGGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCT 1656
Qy 1561 TACTAGCTGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
Db 1657 TACTAGCTGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1716
Qy 1621 ACCCAAGTTCACCAATTTCTGATTTGATCAAGGCCACCATCAAAAGTGAAGTGGCTTC 1680
Db 1717 ACCCAAGTTCACCAATTTCTGATTTGATCAAGGCCACCATCAAAAGTGAAGTGGCTTC 1776
Qy 1681 TAA 1683
Db 1777 TAA 1779

RESULT 5
AAAA4763
ID AAA44763 standard; cDNA; 617 BP.
XX
AC AAA44763;
XX
DT 21-Aug-2000 (first entry)
XX
DE Human secreted expressed sequence tag SEQ ID NO:1338.
XX
KW Human; mouse; chicken; rat; secreted expressed sequence tag; EST;
KW expressed sequence tag; EST; probe; chemokine; chemokine; chemokine;
KW immunomodulatory; haematopoietic; chemokine; chemokine; chemokine;
KW thrombolytic; antiinflammatory; cytosolic; antibacterial; antifungal;
KW antiviral; antidiabetic; antiaslomatic; vulnery; antiparisonian;
KW antilucer; osteoprotective; neuroprotective; nootropic; antiparasitic;
KW cerebroprotective; anticonvulsant; antidepressant; gene therapy;
KW vaccine; autoimmune disorder; multiple sclerosis; allergic condition;
KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
KW central nervous system disorder; Alzheimer's disease; stroke;
KW Parkinson's disease; Huntington's disease; coagulation disorder;
KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease;
KW tumour; infection; depression; psoriasis; ss.
XX
OS Homo sapiens.
XX
PN W0200021991-A1.
XX
PD 20-Apr-2000.

PF 15-OCT-1999: 99MO-US24206.
XX
PR 15-OCT-1998: 98US-0104436.
XX
PA (GEMV) GENETICS INST INC.
XX
PI Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;
PI Metberg D, Treacy M, Bowman MR;
XX
DR WPI: 2000-317938/27.
XX
PT Isolated polynucleotides, and encoded proteins, comprising secreted
PT expressed sequence tags (SESTs), useful for treating various disorders
PT such as autoimmune, infectious, and central nervous system disorders -
XX
PS Claim 1: Page 549: 803pp: English.
XX
CC AAA43426 to AAA45925 represent specifically claimed secreted expressed
CC sequence tags (SESTs), isolated from human, mouse, chicken and rat
CC tissue sources. The SESTs can have a range of activities depending on
CC the tissues they were isolated from. The activities include:
CC chemotactic; proliferative; immunomodulatory; hematopoietic;
CC chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory;
CC cytoskeletal; antibacterial; antifungal; antiviral; antidiabetic;
CC antiasthmatic; vulnereary; antileukemic; osteoprotective;
CC neuroprotective; antiparkinsonian; antipsoriatic; cerebroprotective;
CC anticonvulsant; and antidepressant. The SESTs can be used for gene
CC therapy and in vaccines. The SESTs are useful as probes for the
CC identification and isolation of full-length cDNAs and genomic DNA
CC molecules which correspond to the SESTs. Proteins encoded by the SESTs
CC are useful in assays for determining biological activity and raising
CC antibodies. They may be useful for treatment of autoimmune disorders
CC (multiple sclerosis), insulin dependent diabetes), allergic conditions
CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,
CC osteoporosis, osteoarthritis, central nervous system disorders
CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation
CC disorders (hemophilia, thrombosis), inflammatory disorders (Crohn's
CC disease), tumours, bacterial, fungal or viral infections, depression and
CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given
CC in the exemplification of the present invention.
XX
SQ Sequence 617 BP: 165 A; 169 C; 148 G; 135 T; 0 other;
Query Match 31.3%: Score 526.4; DB 21: Length 617;
Best Local Similarity 99.8%: Pred. No. 9.5e-142;
Matches 527: Conservative 0: Mismatches 1: Indels 0: Gaps 0:

DB 444 GGC CGG GGC CAA TGT CTA TTA CCA CAG ATC CT CCA CTA CCG CTG TCT GTA A CAC 503
QY 421 CCTTACACAGATGTCAGAGTCTCCCAAGTGGTTCTGTATGACAGGACCAACCCCTGCGAG 480
DB 504 CCTTACACAGATGTCAGAGTCTCCCAAGTGGTTCTGTATGACAGGACCAACCCCTGCGAG 563
QY 481 AATGGGGCTACTGCTCCCGCATACCGAGATCCAGATTCACGTTCTG 528
DB 564 AATGGGGCTACTGCTCCCGCATACCGAGATTCACGATTCACCTG 611
RESULT 6
AB199281
ID AB199281 standard. CDNA: 397 BP.
AC
XX AB199281:
XX
XX 07-MAR-2002 (first entry)
XX
XX Mouse ischaemic condition related CDNA sequence SEQ ID NO:123.
XX
XX Mouse: ischaemia; compressive ischaemia; occlusive ischaemia;
XX vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.
XX Mus musculus.
XX W0200188188-A2.
XX
XX 22-NOV-2001.
XX
XX 18-MAY-2001: 2001WO-JP04192.
XX
XX 18-MAY-2000: 2000JP-0145977.
XX
XX (UNIV-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
XX Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX WPI: 2002-034733/04.
XX
XX
XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring
XX expression levels of particular genes defined in the specification or
XX PT by determining the expression profile of a gene group comprising these
XX genes -
XX
PS Claim 2: Page 351-352; 2690pp: English.
XX
XX The present invention describes a method for examining ischaemic
XX conditions, comprising measuring the expression levels of particular
XX genes (1) in a test sample or determining the expression profile of a
XX gene group in the sample comprising genes selected from (1). The method
XX is useful for examining the ischaemic condition (e.g. compressive
XX ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
XX expression levels of particular genes (AB199202 to AB199912, encoding
XX the protein sequences in AB199202 to AB199912) or by determining the
XX expression profile of a gene group comprising these genes. The
XX expression levels or expression profiles produced by these genes are
XX used as an indicator when screening for ischaemic condition-improving
XX drugs or as reagents for ischaemic diseases. AB199913 and AB199914
XX represent PCR primers for a mouse ischaemic condition related sequence,
XX which are used in the exemplification of the present invention.
XX
SQ Sequence 397 BP: 103 A; 105 C; 111 G; 78 T; 0 other;
Query Match 14.3%: Score 241: DB 24: Length 397;
Best Local Similarity 84.4%: Pred. No. 2.6e-59;
Matches 271: Conservative 0: Mismatches 50: Indels 0: Gaps 0:

QY 1363 ACAGAAACAGGAAAGGTCCCGCCAGCTCTGGATGCCAAAGTCAAGCTGATTCGCAC 1422
DB 13 ACAGAAACAGGAAAGGTCCCGCCAGCTCTGGATGCCAAAGTCAAGCTGATTCGCAC 72
QY 1423 ACTTGTGAACCTCCCGCAACTCTATGACCAATGATGATGACAGTATGATCTGCA 1482

```

Db      73 CTTTTTGGCAACTCCGACAGCTCTATGACACACCATGATGACAGATGATTTGTGGC 132
        |||
Qy      1483 GGAATCTTTCAGAAACCTGGCAAGACACCTGCGAGGATGACTGTGAGGCCCTGACC 1542
        |||
Db      133 GGGAGACCTTCAGAAACCCCGGATCAGACACCTGCGAGGATGACTGGGGGCCCTTAACC 192
        |||
Qy      1543 TGTGAGAGGAGCGACCTACTAGCTGTATGAGATGAGCTGGGGCTGGAGTGTGAG 1602
        |||
Db      193 TGTGAGAGGATGAGAACTTACTAGCTGTACGGATTTGTAGCTGGGGCGAGGATGTGG 252
        |||
Qy      1603 AAGAGGCGAGGGGTCTACACCCAGTACCAATTCCTGTAATTTGATCAAAAGCCACCATC 1662
        |||
Db      253 AAGAGCCAGGAGGTCTACACCTACCAAGTTCCTGAATTTGATTAAGACCCACCATG 312
        |||
Qy      1663 AAAAGTGAAGTGGCTTCTAA 1683
        |||
Db      313 CACAGGGAGGCTGGCTCTGTA 333

```

RESULT 7

ABAS8323/c
ID ABAS8323 standard; DNA; 451 BP.

AC ABA58323;

DT 01-FEB-2002 (first entry)

DE Human foetal liver single exon nucleic acid probe #6628.

KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

OS Homo sapiens.

PN W020015727-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00669.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human foetal liver -

XX Claim 1; SEQ ID NO 6628; 639pp + sequence listing; English.

XX The invention relates to a single exon nucleic acid probe for

CC measuring human gene expression in a sample derived from human foetal

CC liver. The single exon nucleic acid probes may be used for predicting,

CC measuring and displaying gene expression in samples derived from human

CC foetal liver. The present sequence is a single exon nucleic acid

CC probe of the invention.

CC Note: The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 451 BP; 116 A; 102 C; 127 G; 106 T; 0 other;

XX Query Match 7.1%; Score 120.2; DB 22; Length 451;

XX Best Local Similarity 97.6%; Pred. No. 2.4e-24;

```

Matches 122; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy      334 TCAGAAAGTCAAAATATGCTGCAGACACCAATGATGGCCGGGGCAATGTCATTAC 383
        |||
Db      451 TCCTACAGTCAAAATATGCTGCAGACCAATGATGGCCGGGGCAATGTCATTAC 392
        |||
Qy      384 CCAGAGTCCTCCCTACTACCGCTGTCTGTAAACACCTTACAGAGTCCACCTGTC 443
        |||
Db      391 CCAGAGTCCTCCCTACTACCGCTGTCTGTAAACACCTTACAGAGTCCACCTGTC 332
        |||
Qy      444 CCAAG 448
        |||
Db      331 CCAAG 327

```

RESULT 8

AAK06420/c
ID AAK06420 standard; DNA; 451 BP.

AC AAK06420;

DT 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe SEQ ID NO: 6411.

KW Human; brain expressed exon; gene expression analysis; probe;

KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;

KW epilepsy; cancer; ss.

OS Homo sapiens.

PN W020015727-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00667.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human

PT brains -

XX Example 4; SEQ ID NO: 6411; 650pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC brain. They can be used to measure gene expression in brain cell samples,

CC which may enable the diagnosis and improved treatment of nervous system

CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

CC epilepsy and cancers. The present sequence is one of the probes of the

CC invention.

XX Sequence 451 BP; 116 A; 102 C; 127 G; 106 T; 0 other;

XX Query Match 7.1%; Score 120.2; DB 22; Length 451;

XX Best Local Similarity 97.6%; Pred. No. 2.4e-24;

XX Matches 122; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

XX Qy 324 TCAGAAAGTCAAAATATGCTGCAGACCAATGATGGCCGGGGCAATGTCATTAC 383

XX Db 451 TCCTACAGTCAAAATATGCTGCAGACCAATGATGGCCGGGGCAATGTCATTAC 392

OY 384 CCAGAGTCTCTCTACTACGCGCTGTGTCTGTAAACACCCCTTACAGAGTCCACGCTCTC 443
|||||
DB 391 CCAGAGTCTCTCTCTACTACGCGCTGTGTCTGTAAACACCCCTTACAGAGTCCACGCTCTC 332
OY 444 CCAAG 448
|||||
DB 331 CCAAG 327

RESULT 9

AAK32096/c
ID AAK32096 standard; DNA: 451 BP.

AAK32096:

06-NOV-2001 (first entry)

Human bone marrow expressed single exon probe SEQ ID NO: 6653.

Human: bone marrow expressed exon; gene expression analysis; probe;
microarray; cancer; leukaemia; lymphoma; myeloma; ss.

Homo sapiens.

WO200157276-A2.

09-AUG-2001.

30-JAN-2001: 2001WO-US00668.

04-FEB-2000: 2000US-0180312.

26-MAY-2000: 2000US-0207456.

30-JUN-2000: 2000US-0608408.

03-AUG-2000: 2000US-0632366.

21-SEP-2000: 2000US-0234687.

27-SEP-2000: 2000US-0236359.

04-OCT-2000: 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Pen n SG, Hanzel DK, Chen W, Rank DR;
WPI: 2001-488900/53.

Human genome-derived single exon nucleic acid probes useful for
analyzing gene expression in human bone marrow -

Example 4: SEQ ID NO: 6653; 658bp + Sequence Listing; English.

The present invention provides a number of single exon nucleic acid
probes which are derived from genomic sequences expressed in the human
bone marrow. They can be used to measure gene expression in bone marrow
samples which may enable the improved diagnosis and treatment of cancers
such as lymphoma, leukaemia and myeloma. The present sequence is one of
the probes of the invention.

Sequence 451 BP: 116 A; 102 C; 127 G; 106 T; 0 other:

Query Match 7.1%; Score 120.2; DB 22; Length 451;
Best Local Similarity 97.6%; Pred. No. 2.4e-24;

Matches 122; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 324 TCAGAAAGTGCAGAAATACGTGCAAGACCAACCATGTGGCGGGCCAAATGTCATTAC 383
|||

DB 451 TCCTACAGTGCAGAAATACGTGCAAGACCAACCATGTGGCGGGCCAAATGTCATTAC 392
|||||

OY 384 CCAGAGTCTCTCTACTACGCGCTGTGTCTGTAAACACCCCTTACAGAGTCCACGCTCTC 443
|||||

DB 391 CCAGAGTCTCTCTACTACGCGCTGTGTCTGTAAACACCCCTTACAGAGTCCACGCTCTC 332
|||||

OY 444 CCAAG 448
|||||

DB 331 CCAAG 327

RESULT 10

AA137947/c
ID AA137947 standard; DNA: 451 BP.

AA137947:

17-OCT-2001 (first entry)

Probe #6633 used to measure gene expression in human placenta sample.

Probe: microarray; human; placenta; antenatal diagnosis;

genetic disorder; ss.

Homo sapiens.

WO200157272-A2.

09-AUG-2001.

30-JAN-2001: 2001WO-US00663.

04-FEB-2000: 2000US-0180312.

26-MAY-2000: 2000US-0207456.

30-JUN-2000: 2000US-0608408.

03-AUG-2000: 2000US-0632366.

21-SEP-2000: 2000US-0234687.

27-SEP-2000: 2000US-0236359.

04-OCT-2000: 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Pen n SG, Hanzel DK, Chen W, Rank DR;
WPI: 2001-488897/53.

Human genome-derived single exon nucleic acid probes useful for
analyzing gene expression in human placenta -

Claim 25: SEQ ID NO 6633; 654bp; English.

Sequence 451 BP: 116 A; 102 C; 127 G; 106 T; 0 other:

Query Match 7.1%; Score 120.2; DB 22; Length 451;
Best Local Similarity 97.6%; Pred. No. 2.4e-24;

Matches 122; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 324 TCAGAAAGTGCAGAAATACGTGCAAGACCAACCATGTGGCGGGCCAAATGTCATTAC 383
|||

DB 451 TCCTACAGTGCAGAAATACGTGCAAGACCAACCATGTGGCGGGCCAAATGTCATTAC 392
|||||

OY 384 CCAGAGTCTCTCTACTACGCGCTGTGTCTGTAAACACCCCTTACAGAGTCCACGCTCTC 443
|||||

DB 391 CCAGAGTCTCTCTACTACGCGCTGTGTCTGTAAACACCCCTTACAGAGTCCACGCTCTC 332
|||||

OY 444 CCAAG 448
|||||

DB 331 CCAAG 327

RESULT 11

ABS06863/c
ID ABS06863 standard; DNA: 451 BP.

ABS06863:

XX 19-AUG-2002 (first entry)
 DT
 XX
 DE Human genome-derived single exon probe from lung SEQ ID No 6854.
 XX
 KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberosus sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200186003-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00665.
 XX
 PR 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 DR WPI: 2002-114183/15.
 XX
 PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 XX
 PS Claim 1; SEQ ID No 6854; 634bp; English.

CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karsenger syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a single exon
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp://ipo.int/pub/published_pcl_sequences.
 XX
 SQ Sequence 451 BP; 116 A; 102 C; 127 G; 106 T; 0 other;
 Query Match 7.1%; Score 120.2; DB 24; Length 451;
 Best Local Similarity 97.6%; Pred. No. 2.4e-24;
 Matches 122; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 OY 324 TCAGAAAGTCGCAAAATACGTGCAAGGACCAACCATGTGCGGCGCCATGTCTATTAC 383
 DB 451 TCCTACAGTGCAGAAATACGTGCAAGGACCAACCATGTGCGGCGCCATGTCTATTAC 392
 OY 384 CCAGAGTCCCTCCCTACTACGCGTGTGTCTTAACACCCCTTACACAGGATCCAGAGTCGTC 443
 DB 391 CCAGAGTCCCTCCCTACTACGCGTGTGTCTTAACACCCCTTACACAGGATCCAGAGTCGTC 332
 OY 444 CCAAG 448
 DB 331 CCAAG 327
 RESULT 12
 ABA70919/C
 ID ABA70919 standard; DNA; 117 BP.
 XX
 AC ABA70919;
 XX
 DT 01-FEB-2002 (first entry)
 XX
 DE Human foetal liver single exon nucleic acid probe #19224.
 XX
 KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157277-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00669.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 DR WPI: 2001-483447/52.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human foetal liver -
 XX
 PS Claim 4; SEQ ID NO 19224; 639bp + sequence listing; English.

CC The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC foetal liver. The present sequence is a single exon nucleic acid


```

RESULT 15
AA151106/C
ID AA151106 standard; DNA; 117 bp.
XX
AC AA151106;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #19792 used to measure gene expression in human placenta sample.
XX
KM Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00663.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta.
XX
PS Claim 25; SEQ ID No 19792; 654bp; English.
XX
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders.
XX
SQ Sequence 117 bp; 24 A; 24 C; 40 G; 29 T; 0 other;

Query Match          7.0%; Score 117; DB 22; Length 117;
Best Local Similarity 100.0%; Pred. No. 1e-23;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 332 TGCATAATACGTGCAAGACACCATGTGGCGGGCCATGTCATATTACCGAGATC 391
DB 117 TGCATAATACGTGCAAGACACCATGTGGCGGGCCATGTCATATTACCGAGATC 58
OY 392 CTCCTACTACCGCTGTCTGTAAACACCTTACACAGGTCACAGCTCTCCCAAG 448
DB 57 CTCCTACTACCGCTGTCTGTAAACACCTTACACAGGTCACAGCTCTCCCAAG 1

Search completed: March 6, 2003, 16:01:40
Job time : 348.5 secs

```